PRINCIPIOS de ANÁLISIS MATEMÁTICO

WALTER. RUDIN

TALLERES
ESTUDIANTILES
CIENCIAS
UNAM

$$g'(x) = 1 + \left\{2x - \frac{2i}{x}\right\} e^{i/x^2} \quad (0 < x < 1),$$

$$|g'(x)| \geq \left|2x - \frac{2i}{x}\right| - 1 \geq \frac{2}{x} - 1.$$

$$\left|\frac{f'(x)}{g'(x)}\right| = \frac{1}{|g'(x)|} \leq \frac{x}{2-x}$$

$$\lim_{x \to 0} \frac{f(x)}{g(x)} = 0.$$

Traducido por:
Miguel Irán Alcerreca
Revisado por:
Luis Briseño Aguirre

Edición impresa:
McGraw Hill

3a Edición 1980

PRINCIPIOS DE ANÁLISIS MATEMÁTICO

PRINCIPIOS DE ANALISIS MATEMATICO, 3a. edición

Atlacomulco 499-501, Naucalpan de Juárez, Edo. de México
Miembro de la Cámara Nacional de la Industria Editorial, Reg. Núm. 465

ISBN 968-6046-82-8

Traducido de la tercera edición en inglés de
PRINCIPLES OF MATHEMATICAL ANALYSIS

ISBN-0-07-054235-X

1 2 3 4 5 6 7 8 9 0 IMPRO-80 8 1 2 3 4 5 6 7 9 0

Impreso en México Printed in Mexico

Esta obra se terminó de imprimir en julio de 1980
en Impresora Roma, S. A.
Tomás Vázquez 152, Col. Ampliación Moderna,
México 13, D. F.

Se tiraron 6 000 ejemplares

Principios de Análisis Matemático

TERCERA EDICIÓN

WALTER RUDIN
Profesor de Matemáticas
University of Wisconsin-Madison

Traducido por:
LIC. MIGUEL IRÁN ALCERRECA SÁNCHEZ
Licenciado en Física y Matemáticas-
Instituto Politécnico Nacional,
Candidato a Maestro en Ciencias en
Ingeniería Nuclear,
Profesor Asociado, Escuela Superior de
Ingeniería Mecánica y Eléctrica del
Instituto Politécnico Nacional,
México

Revisado por:
LUIS BRISEÑO AGUIRRE
Facultad de Ciencias,
Departamento de Matemáticas en la
Universidad Nacional Autónoma de México,
México

LIBROS McGRAW-HILL

México • Bogotá • Madrid • Panamá • San Juan • Sao Paulo • Nueva York
Auckland • Guatemala • Hamburgo • Johannesburg • Lisboa • Londres
Montreal • Nueva Delhí • París • San Francisco
Singapur • St. Louis • Sydney • Tokio • Toronto

CONTENIDO

Prefacio ix

Capítulo 1 **Los Sistemas de los Números Reales y de los Complejos** **1**

Introducción 1
Conjuntos Ordenados 3
Campos 5
El Campo Real 9
El Sistema Extendido de los Números Reales 12
El Campo Complejo 13
Espacios Euclidianos 17
Apéndice 18
Ejercicios 23

Capítulo 2 **Topología Básica** **26**

Conjuntos Finitos, Numerables y No Numerables 26
Espacios Métricos 33
Conjuntos Compactos 39
Conjuntos Perfectos 44

	Conjuntos Conexos	46
	Ejercicios	46
Capítulo 3	**Sucesiones Numéricas y Series**	**50**
	Sucesiones Convergentes	50
	Subsucesiones	54
	Sucesiones de Cauchy	55
	Límites Superior e Inferior	59
	Algunas Sucesiones Especiales	61
	Series	62
	Series de Términos No Negativos	64
	El Número *e*	67
	Criterios de la Raíz y de la Razón	70
	Series de Potencias	73
	Suma por Partes	74
	Convergencia Absoluta	76
	Adición y Multiplicación de Series	77
	Reordenamientos	80
	Ejercicios	83
Capítulo 4	**Continuidad**	**89**
	Límites de Funciones	89
	Funciones Continuas	91
	Continuidad y Compacticidad	95
	Continuidad y Conexibilidad	100
	Discontinuidades	100
	Funciones Monótonas	102
	Límites Infinitos y Límites en el Infinito	104
	Ejercicios	105
Capítulo 5	**Diferenciación**	**110**
	Derivada de una Función Real	110
	Teoremas del Valor Medio	114
	Continuidad de las Derivadas	115
	Regla de L'Hospital	116
	Derivadas de Orden Superior	118
	Teorema de Taylor	118
	Diferenciación de Funciones Vectoriales	119
	Ejercicios	121
Capítulo 6	**La Integral de Riemann-Stieltjes**	**129**
	Definición y Existencia de la Integral	129
	Propiedades de la Integral	137
	Integración y Diferenciación	143
	Integración de Funciones Vectoriales	145
	Curvas Rectificables	146
	Ejercicios	148

Capítulo 7	**Sucesiones y Series de Funciones**	**153**
	Discusión del Problema Principal	154
	Convergencia Uniforme	157
	Convergencia Uniforme y Continuidad	159
	Convergencia Uniforme e Integración	162
	Convergencia Uniforme y Diferenciación	163
	Familias Equicontinuas de Funciones	165
	Teorema de Stone-Weierstrass	170
	Ejercicios	177
Capítulo 8	**Algunas Funciones Especiales**	**184**
	Series de Potencias	184
	Las Funciones Exponencial y Logarítmica	191
	Funciones Trigonométricas	195
	La Completitud Algebraica del Campo Complejo	198
	Series de Fourier	199
	La Función Gamma	206
	Ejercicios	211
Capítulo 9	**Funciones de Varias Variables**	**219**
	Transformaciones Lineales	219
	Diferenciación	227
	El Principio de la Contracción	237
	Teorema de la Función Inversa	238
	El Teorema de la Función Implícita	241
	El Teorema del Rango	246
	Determinantes	250
	Derivadas de Orden Superior	254
	Diferenciación de Integrales	256
	Ejercicios	258
Capítulo 10	**Integración de Formas Diferenciales**	**265**
	Integración	265
	Mapeos Primitivos	268
	Particiones de la Unidad	271
	Cambio de Variable	272
	Formas Diferenciales	274
	Cadenas y Símplex	288
	Teorema de Stokes	295
	Formas Cerradas y Exactas	298
	Análisis Vectorial	304
	Ejercicios	313
Capítulo 11	**Teoría de Lebesgue**	**324**
	Funciones de Conjuntos	324
	Construcción de la Medida de Lebesgue	327

Espacios de Medida 335
Funciones Medibles 335
Funciones Simples 338
Integración 339
Comparación con la Integral de Riemann 348
Integración de Funciones Complejas 351
Funciones de Clase $\mathscr{L}^2$ 352
Ejercicios 359

Bibliografía 363

Lista de Símbolos Especiales 365

Índice 367

PREFACIO

Se pretende que este libro sirva como texto para el curso de análisis que reciben normalmente los estudiantes avanzados de licenciatura o para los estudiantes graduados de primer año que estudian Matemáticas.

La presente edición con algo más de material, un poco de menos omisiones, y un reordenamiento considerable, cubre en esencia los mismos temas que la segunda. Espero que dichos cambios hagan más accesible y atractivo el material a los estudiantes que reciben tal curso.

La experiencia me ha convencido que, pedagógicamente, es erróneo (aunque desde el punto de vista lógico es correcto) comenzar con la construcción de los números reales a partir de los racionales. Simplemente en un principio, la mayoría de los estudiantes no aprecian la necesidad de hacerlo. Por esto se presenta el sistema de los números como un campo que posee la propiedad de la mínima cota superior, y se efectúan rápidamente algunas aplicaciones interesantes de esta propiedad. Sin embargo no se omite la construcción de Dedekind. Ahora se encuentra en el apéndice del capítulo 1, en donde puede estudiarse y deleitarse siempre y cuando se tenga la madurez adecuada.

Se volvió a escribir casi todo el material sobre funciones de varias variables, completándolo con muchos detalles y más motivación con ejemplos. La demostración del teorema de la función inversa, que es el tema clave del capítulo 9, se simplifica con el teorema de punto fijo referente a mapeos de

contracción. Las formas diferenciales se estudian con más detalle. Se incluyen además varias aplicaciones del teorema de Stokes.

En lo que se refiere a cambios, el capítulo sobre la integral de Riemann-Stieltjes se ha equilibrado un poco; también se adicionó al capítulo 8 una pequeña sección sobre la función gama para que el lector la desarrolle, y hay un número bastante grande de ejercicios nuevos, la mayoría de ellos con sugerencias bastante detalladas.

También incluí varias referencias sobre artículos publicados en el *American Mathematical Monthly* y el *Mathematics Magazine*, esperando que a los estudiantes se les desarrolle el hábito de consultar las publicaciones científicas. R. B. Burckel muy amablemente me proporcionó la mayoría de las referencias.

Numerosas personas, tanto estudiantes como maestros, durante muchos años me han enviado correcciones, críticas y comentarios acerca de las ediciones anteriores de este libro. Las aprecio y aprovecho la oportunidad para expresarles mis más sinceros agradecimientos a todos los que me han escrito.

WALTER RUDIN

1

LOS SISTEMAS DE LOS NÚMEROS REALES Y DE LOS COMPLEJOS

INTRODUCCIÓN

Para que una presentación de los conceptos principales del análisis (tales como la convergencia, continuidad, diferenciación y la integración) sea satisfactoria, debe basarse en el concepto de número definido con exactitud. Sin embargo, aquí no se abordará el tema sobre los axiomas que gobiernan la aritmética de los enteros, pero se supondrá que se conocen bien los números racionales (es decir, los de la forma m/n, en donde m y n son enteros y $n \neq 0$).

El sistema de los números racionales visto como un campo o un conjunto ordenado, es inadecuado para muchos casos. (Los conceptos de campo y conjunto ordenado se definirán en las Secs. 1.6 y 1.12.) Por ejemplo, no existe un racional p tal que $p^2 = 2$. (Se demostrará esto a continuación.) Lo anterior conduce a la introducción de los llamados "números irracionales", que con frecuencia se representan como desarrollos decimales infinitos y se consideran "aproximados" por los decimales finitos correspondientes. Entonces la sucesión

$$1, 1.4, 1.41, 1.414, 1.4142, \ldots$$

"tiende a $\sqrt{2}$". Puede surgir la siguiente pregunta, a menos que el número irracional $\sqrt{2}$ esté evidentemente definido: ¿Qué quiere decir que la sucesión "tienda a"?

Esta clase de pregunta se puede contestar tan pronto como se construya el llamado "sistema de los números reales".

1.1 Ejemplo Empecemos demostrando que la ecuación

$$p^2 = 2 \tag{1}$$

no se puede satisfacer por ningún número racional p. Para ello supongamos que se satisface: podremos escribir $p = m/n$, donde m y n son enteros y además podremos elegirlos de modo que los dos no sean pares. Supongamos que lo hemos hecho; entonces (1) implica que

$$m^2 = 2n^2, \tag{2}$$

Lo que muestra que m^2 es par. Por tanto, m es par (si m fuera impar, m^2 también lo sería) y m^2 es divisible por 4. De aquí se deduce que el segundo miembro de (2) es divisible por 4, y por tanto n^2 es par, lo que implica que lo sea n.

Por consiguiente, el suponer que se verifica (1) nos lleva a la conclusión que m y n son los dos pares, en contra de la elección que para ellos habíamos hecho. Luego (1) es imposible para un número racional p.

Examinemos ahora la situación algo más de cerca. Sea A el conjunto de todos los números racionales positivos p, tales que $p^2 < 2$, y B el de todos aquellos para los que $p^2 > 2$. Veremos que *A no contiene ningún número que sea mayor, ni B ninguno que sea menor, que todos los demás.*

Más explícitamente, para cada p de A podemos hallar otro número racional q, también en A, tal que $p < q$; y para cada p de B otro q de B tal que $q < p$.

Para esto, se asocia a cada número racional $p > 0$ el número

$$q = p - \frac{p^2 - 2}{p + 2} = \frac{2p + 2}{p + 2}. \tag{3}$$

Por lo que

$$q^2 - 2 = \frac{2(p^2 - 2)}{(p + 2)^2}. \tag{4}$$

Si p pertenece a A, entonces $p^2 - 2 < 0$, de (3) se puede ver que $q > p$ y de (4) $q^2 < 2$. Por lo que q pertenece a A.

Si p pertenece a B, entonces $p^2 - 2 > 0$, de (3) se puede ver que $0 < q < p$ y de (4) $q^2 > 2$. Por lo que q pertenece a B.

1.2 Observación El objeto de la discusión anterior ha sido demostrar que el sistema de los números racionales tiene ciertas lagunas, a pesar del hecho

de que entre dos números racionales, siempre hay otro: Si $r < s$ entonces $r < (r + s)/2 < s$. El sistema de los números reales llena estas lagunas, y es la razón principal del papel tan fundamental que desempeña este sistema en el análisis.

Para dilucidar su estructura y la de los números complejos, se empezará con una discusión breve de los conceptos generales de *conjunto ordenado* y *campo*.

A continuación se da parte de la terminología estándar de la teoría de conjuntos que se usará en todo este libro.

1.3 Definiciones Si A es un conjunto cualquiera (cuyos elementos pueden ser números u objetos cualesquiera), se escribirá $x \in A$ para expresar que x es un miembro (o elemento) de A.

Si x no es un miembro de A, se escribirá: $x \notin A$.

El conjunto que no contiene ningún elemento se llamará *conjunto vacío*. Si un conjunto tiene al menos un elemento, es un conjunto *no vacío*.

Si A y B son conjuntos y todo elemento de A es un elemento de B, se dice entonces que A es un subconjunto de B, y se escribe $A \subset B$, o $B \supset A$. Si además existe un elemento de B que no pertenece a A, entonces A será un subconjunto *propio* de B. Obsérvese que $A \subset A$ para todo conjunto A.

Si $A \subset B$ y $B \subset A$, se escribe $A = B$. De otra manera será $A \neq B$.

1.4 Definición En todo el capítulo 1, el conjunto de los números racionales se representará con una Q.

CONJUNTOS ORDENADOS

1.5 Definición Si S es un conjunto, un *orden* en S es una relación representada por el símbolo $<$ y tiene las dos propiedades siguientes:

(*i*) Si $x \in S$ y $y \in S$, una y sólo una de las proposiciones siguientes es cierta:

$$x < y, \qquad x = y, \qquad y < x$$

(*ii*) Si $x, y, z \in S$, y si $x < y$ y $y < z$, entonces $x < z$.

La proposición "$x < y$" se lee "x es menor que y" o "x es más pequeño que y" o también "x precede a y".

Con frecuencia conviene escribir $y > x$ en vez de $x < y$.

La notación $x \leq y$ indica que $x < y$ o $x = y$, sin especificar cuál de las dos se cumple. Dicho de otra manera, $x \leq y$ es la negación de $x > y$.

1.6 Definición Un *conjunto ordenado* es aquel en el que se ha definido un orden.

Por ejemplo, Q es un conjunto ordenado si se define $r < s$ si $s - r$ es un número racional positivo.

1.7 Definición Supóngase que S es un conjunto ordenado, y $E \subset S$. Se dice que E es un conjunto *acotado superiormente* si existe un $\beta \in S$ tal que $x \leq \beta$ para cada $x \in E$, y a β se le denomina la *cota superior* de E.

De la misma manera se definen las cotas inferiores (con $\geq$ en lugar de $\leq$).

1.8 Definición Supóngase que S es un conjunto ordenado, $E \subset S$, y E está acotado superiormente. Supóngase, además, que existe un $\alpha \in S$ con las siguientes propiedades:

(*i*) α es una cota superior de E.
(*ii*) Si $\gamma < \alpha$, entonces γ no es una cota superior de E.

Entonces α se denomina la *mínima cota superior* de E [de (*ii*) es evidente que a lo más hay una α] o el *supremum* de E, y se escribe

$$\alpha = \sup E.$$

La *máxima cota inferior* o *infimum* de un conjunto E que está acotado inferiormente, se define de la misma manera: La proposición

$$\alpha = \inf E$$

significa que α es una cota inferior de E y que ninguna β con $\beta > \alpha$ es una cota inferior de E.

1.9 Ejemplos

(*a*) Considérense los conjuntos A y B del Ejemplo 1.1 como subconjuntos del conjunto ordenado Q. El conjunto A está acotado superiormente. De hecho, las cotas superiores de A son los miembros de B. Como B no contiene un miembro más pequeño, entonces *A no tiene mínima cota superior en Q.*

De igual manera, B está acotado inferiormente: el conjunto de todas las cotas inferiores de B consta de A y de todos los $r \in Q$ con $r \leq 0$. Como A no tiene un miembro más grande, entonces *B no tiene máxima cota inferior en Q.*

(*b*) Si existe $\alpha = \sup E$, puede o no, ser α miembro de E. Sea por ejemplo, E_1 el conjunto de todos los $r \in Q$ con $r < 0$. Y E_2 el conjunto de todos los $r \in Q$ con $r \leq 0$. Entonces

$$\sup E_1 = \sup E_2 = 0,$$

y $0 \notin E_1$, $0 \in E_2$.
(c) Si E consta de todos los números $1/n$, en donde $n = 1, 2, 3, \ldots$. Entonces $\sup E = 1$ y pertenece a E, el $\inf E = 0$, y no pertenece a E.

1.10 Definición Se dice que un conjunto ordenado S tiene la *propiedad de la mínima cota superior* si lo siguiente es cierto:

Si $E \subset S$, E no es vacío y está acotado superiormente, entonces E existe en S.

El ejemplo 1.9(a) muestra que Q no tiene la propiedad de la mínima cota superior.

Ahora se mostrará que hay bastante relación entre las máximas cotas inferiores y las mínimas cotas superiores, y que todo conjunto ordenado que tenga la propiedad de la mínima cota superior tiene también la de la máxima cota inferior.

1.11 Teorema *Supóngase que S es un conjunto ordenado con la propiedad de la mínima cota superior, $B \subset S$ y B es no vacío y acotado inferiormente. Ahora sea L el conjunto de todas las cotas inferiores de B. Entonces*

$$\alpha = \sup L$$

existe en S y $\alpha = \inf B$.

En particular, el $\inf B$ *existe en S.*

Demostración Como B está acotado inferiormente, L no es vacío. Ya que L consta exactamente de los $y \in S$ que además satisfacen la desigualdad $y \leq x$ para cada $x \in B$, es evidente que *cada $x \in B$ es una cota superior de L.* De aquí que L está acotado superiormente. La hipótesis referente a S implica por tanto, que L tiene un supremum en S; llamémosle α.

Si $\gamma < \alpha$ entonces (véase la Definición 1.8) γ no es cota superior de L, de aquí que $\gamma \notin B$. Y resulta que $\alpha \leq x$ para cada $x \in B$. Entonces $\alpha \in L$.

Si $\alpha < \beta$ se tiene que $\beta \notin L$ debido a que α es cota superior de L.

Se ha mostrado que $\alpha \in L$, pero que $\beta \notin L$ si $\beta > \alpha$. En otras palabras, α es una cota inferior de B, pero β no lo es si $\beta > \alpha$. Esto significa que $\alpha = \inf B$.

CAMPOS

1.12 Definición Un *campo* es un conjunto F con dos operaciones, llamadas *adición* y *multiplicación*, que satisfacen los axiomas siguientes, (A), (M) y (D) llamados "axiomas de campo":

(A) Axiomas para la adición

(A1) Si $x \in F$ y $y \in F$, entonces su suma, $x + y$ está en F.
(A2) La adición es conmutativa: $x + y = y + x$ para toda $x, y \in F$.
(A3) La adición es asociativa: $(x + y) + z = x + (y + z)$ para toda $x, y, z \in F$.
(A4) F contiene un elemento 0 tal que $0 + x = x$ para cada $x \in F$.
(A5) A cada $x \in F$ le corresponde un elemento $-x \in F$ tal que

$$x + (-x) = 0.$$

(M) Axiomas para la multiplicación

(M1) Si $x \in F$ y $y \in F$, entonces su producto xy está en F.
(M2) La multiplicación es conmutativa: $xy = yx$ para toda $x, y \in F$.
(M3) La multiplicación es asociativa: $(xy)z = x(yz)$ para toda $x, y, z \in F$.
(M4) F contiene un elemento $1 \neq 0$ tal que $1x = 0$ para cada $x \in F$.
(M5) Si $x \in F$ y $x \neq 0$, entonces existe un elemento $1/x \in F$ tal que

$$x \cdot (1/x) = 1.$$

(D) La ley distributiva

$$x(y + z) = xy + xz$$

se cumple para toda $x, y, z \in F$.

1.13 Observaciones

(*a*) Se acostumbra escribir (en cualquier campo)

$$x - y, \frac{x}{y}, x + y + z, xyz, x^2, x^3, 2x, 3x, \ldots$$

en vez de

$$x + (-y), x \cdot \left(\frac{1}{y}\right), (x + y) + z, (xy)z, xx, xxx, x + x, x + x + x, \ldots.$$

(*b*) Los axiomas de campo evidentemente se cumplen en el conjunto Q de todos los números racionales, siempre y cuando la adición y la multiplicación signifiquen lo acostumbrado. Por ende, Q es un campo.
(*c*) Aunque nuestro propósito no es el de estudiar campos (o cualquier otra estructura algebraica) con mucho detalle, vale la pena de-

mostrar que algunas propiedades conocidas de Q son consecuencia de los axiomas de campo; una vez hecho esto, no se necesitará volver a hacerlo para los números reales o para los complejos.

1.14 Proposición *Los axiomas para la adición implican las siguientes proposiciones:*

(*a*) *Si* $x + y = x + z$, *entonces* $y = z$.
(*b*) *Si* $x + y = x$, *entonces* $y = 0$.
(*c*) *Si* $x + y = 0$, *entonces* $y = -x$.
(*d*) $-(-x) = x$.

La proposición (*a*) es la ley de la cancelación. Nótese que (*b*) asegura la unicidad del elemento cuya existencia se supuso en (A4), y que (*c*) hace lo mismo para (A5).

Demostración Si $x + y = x + z$, de los axiomas (A) se tiene

$$\begin{aligned} y = 0 + y = (-x + x) + y &= -x + (x + y) \\ &= -x + (x + z) = (-x + x) + z = 0 + z = z. \end{aligned}$$

Esto demuestra (*a*). Tómese $z = 0$ en (*a*) para obtener (*b*), y $z = -x$ en (*a*) para obtener (*c*).
Como $-x + x = 0$, (*c*) (con $-x$ en lugar de x) esto produce (*d*).

1.15 Proposición *Los axiomas para la multiplicación implican las proposiciones siguientes:*

(*a*) *Si* $x \neq 0$ *y* $xy = xz$, *entonces* $y = z$.
(*b*) *Si* $x \neq 0$ *y* $xy = x$, *entonces* $y = 1$.
(*c*) *Si* $x \neq 0$ *y* $xy = 1$, *entonces* $y = 1/x$.
(*d*) *Si* $x \neq 0$, *entonces* $1/(1/x) = x$.

La demostración es similar a la de la proposición 1.14 y por eso no se da aquí.

1.16 Proposición *Los axiomas de campo implican las siguientes proposiciones, para cualquier* $x, y, z \in F$:

(*a*) $0x = 0$.
(*b*) *Si* $x \neq 0$ *y* $y \neq 0$, *entonces* $xy \neq 0$.
(*c*) $(-x)y = -(xy) = x(-y)$.
(*d*) $(-x)(-y) = xy$.

Demostración $0x + 0x = (0 + 0)x = 0x$. En consecuencia, 1.14(*b*) implica que $0x = 0$, y por esto se cumple (*a*).

Supóngase ahora $x \neq 0$, $y \neq 0$, pero que $xy = 0$. Entonces de (a) se tiene

$$1 = \left(\frac{1}{y}\right)\left(\frac{1}{x}\right)xy = \left(\frac{1}{y}\right)\left(\frac{1}{x}\right)0 = 0,$$

una contradicción. Por lo tanto se cumple (b).

La primera igualdad en (c) proviene de

$$(-x)y + xy = (-x + x)y = 0y = 0,$$

combinada con 1.14(c); la otra mitad de (c) se demuestra de la misma forma. Finalmente de (c) y 1.14(d) se obtiene

$$(-x)(-y) = -[x(-y)] = -[-(xy)] = xy$$

1.17 Definición Un *campo ordenado* es un *campo F* que a su vez es un *conjunto ordenado*, y que tiene las siguientes propiedades:

(i) $x + y < x + z$ si $x, y, z \in F$ y $y < z$.
(ii) $xy > 0$ si $x \in F$, $y \in F$, $x > 0$, y $y > 0$.

Si $x > 0$, se dice que x es *positivo*; si $x < 0$, entonces es *negativo*.

Por ejemplo, Q es un campo ordenado.

En cada campo ordenado se aplican todas las reglas conocidas de las desigualdades: la multiplicación por cantidades positivas (negativas) preserva (invierte) las desigualdades, ningún cuadrado es negativo, etcétera. En la siguiente proposición se listan algunas de estas.

1.18 Proposición *En todo campo ordenado, las siguientes proposiciones son verdaderas*:

(a) *Si* $x > 0$, *entonces* $-x < 0$, *y viceversa.*
(b) *Si* $x > 0$ *y* $y < z$, *entonces* $xy < xz$.
(c) *Si* $x < 0$ *y* $y < z$, *entonces* $xy > xz$.
(d) *Si* $x \neq 0$, *entonces* $x^2 > 0$. *En particular*, $1 > 0$.
(e) *Si* $0 < x < y$, *entonces* $0 < 1/y < 1/x$.

Demostración

(a) Si $x = 0$, entonces $0 = -x + x > -x + 0$, así que $-x < 0$. Si $x < 0$, entonces $0 = -x + x < -x + 0$, de manera que $-x > 0$. Esto demuestra (a).

(b) Debido a que $z > y$, se tiene $z - y > y - y = 0$, por consiguiente $x(z-y) > 0$, y por tanto

$$xz = x(z - y) + xy > 0 + xy = xy.$$

(c) De (a), (b) y la Proposición 1.16(c) se obtiene,

$$-[x(z-y)] = (-x)(z-y) > 0,$$

de aquí que $x(z - y) < 0$, y en consecuencia $xz < xy$.
(d) Si $x > 0$, la parte (ii) de la Definición 1.17 nos da $x^2 > 0$. Si $x < 0$, entonces $-x > 0$, por ende $(-x)^2 > 0$. Pero $x^2 = (-x)^2$, debido a la Proposición 1.16(d). Ya que $1 = 1^2$, $1 > 0$.
(e) Si $y > 0$ y $v \leq 0$, entonces $yv \leq 0$. Pero $y \cdot (1/y) = 1 > 0$. En consecuencia, $1/y > 0$. De la misma forma, $1/x > 0$. Si se multiplican ambos miembros de la desigualdad $x < y$, por la cantidad positiva $(1/x)(1/y)$, se obtiene $1/y < 1/x$.

EL CAMPO REAL

Ahora se establecerá el *teorema de existencia* que es la base de este capítulo.

1.19 Teorema *Existe un campo ordenado R con la propiedad de la mínima cota superior.*
Además, R contiene como subcampo a Q.

La segunda proposición significa que $Q \subset R$ y que cuando se aplican las operaciones de adición y multiplicación en R a los miembros de Q, éstas coinciden con las operaciones más comunes en los números racionales; también los números racionales positivos son elementos positivos de R.

A los miembros de R se les denomina *números reales*.

Debido a que la demostración del Teorema 1.19 es bastante larga y un poco tediosa, se presenta en el Apéndice del capítulo 1. En realidad, la demostración construye R a partir de Q.

El siguiente teorema podría obtenerse a partir de esta construcción con un poco de esfuerzo adicional. No obstante, es preferible obtenerlo a partir del Teorema 1.19, debido a que éste proporciona una buena ilustración de lo que se puede hacer con la propiedad de la mínima cota superior.

1.20 Teorema

(a) *Si $x \in R$, $y \in R$ y $x > 0$, entonces hay un entero positivo n tal que*

$$nx > y.$$

(b) *Si $x \in R$, $y \in R$, y $x < y$, entonces existe un $p \in Q$ tal que*

$$x < p < y.$$

La parte (a) se conoce comúnmente como la *propiedad arquimediana* de R. La parte (b) significa que Q es *denso* en R: es decir, entre dos números reales cualesquiera hay un racional.

Demostración

(a) Sea A el conjunto de todos los nx, en donde n varía en los enteros positivos. Si (a) no fuera cierto, entonces y sería una cota superior de A, y A entonces tiene una *mínima* cota superior en R. Hagamos $\alpha = \sup A$. Ya que $x > 0$, $\alpha - x < \alpha$, y $\alpha - x$ no es una cota superior de A. Por consiguiente, $\alpha - x < mx$ para algún entero positivo n. De donde $\alpha < (m + 1)x \in A$, lo cual es imposible, porque α es una cota superior de A.

(b) Debido a que $x < y$, se tiene $y - x > 0$, y (a) proporciona un entero positivo n tal que

$$n(y - x) > 1.$$

Aplicando (a) nuevamente, se obtienen los enteros positivos m_1 y m_2 de tal manera que $m_1 > nx$, $m_2 > -nx$. Entonces

$$-m_2 < nx < m_1.$$

En consecuencia hay un entero m (con $-m_2 \leq m \leq m_1$) tal que

$$m - 1 \leq nx < m.$$

Si estas desigualdades se combinan, se obtiene

$$nx < m \leq 1 + nx < ny.$$

Como $n > 0$, se deduce que

$$x < \frac{m}{n} < y.$$

Esto demuestra (b), con $p = m/n$.

Ahora se demostrará la existencia de las raíces n-ésimas de los reales positivos. Esta prueba mostrará cómo puede manejarse en R, la dificultad mencionada en la Introducción (la irracionalidad de $\sqrt{2}$).

1.21 Teorema *Para todo número real $x > 0$ y cada entero $n > 0$ hay un número real $y > 0$, y uno solo, tal que $y^n = x$.*

Este número y se escribe $\sqrt[n]{x}$, o $x^{1/n}$.

Demostración Que como máximo hay un y con tal propiedad, es evidente, pues, $0 < y_1 < y_2$ implica $y_1^n < y_2^n$.

Sea E el conjunto formado por todos los números reales positivos t, tales que $t^n < x$.

Si $t = x/(1 + x)$, será $0 < t < 1$; por tanto, $t^n \leq t < x$, por lo que $t \in E$ y E no es vacío.

Si $t > 1 + x$, entonces $t^n > t > x$, así que $t \notin E$. Es decir, $1 + x$ es una cota superior de E.

El Teorema 1.19 implica por lo tanto, la existencia de

$$y = \sup E.$$

Para demostrar que $y^n = x$ se mostrará que cada una de las desigualdades $y^n < x$ y $y^n > x$ conduce a una contradicción.

La identidad $b^n - a^n = (b - a)(b^{n-1} + b^{n-2}a + \cdots + a^{n-1})$ produce la desigualdad

$$b^n - a^n < (b - a)nb^{n-1}$$

cuando $0 < a < b$.

Ahora supóngase que $y^n < x$. Si se escoge h de tal forma que $0 < h < 1$ y

$$h < \frac{x - y^n}{n(y + 1)^{n-1}}.$$

Si se hace $a = y$, $b = y + h$. Entonces

$$(y + h)^n - y^n < hn(y + h)^{n-1} < hn(y + 1)^{n-1} < x - y^n.$$

Es decir $(y + h)^n < x$, y $y + h \in E$. Ya que $y + h > y$, esto contradice el hecho de que y es una cota superior de E.

Suponiendo que $y^n > x$. Hágase

$$k = \frac{y^n - x}{ny^{n-1}}.$$

Entonces $0 < k < y$. Si $t \geq y - k$, se concluye que

$$y^n - t^n \leq y^n - (y - k)^n < kny^{n-1} = y^n - x.$$

Por lo que $t^n > x$, y $t \notin E$. Se deduce que $y - k$ es una cota superior de E.

Pero $y - k < y$, esto contradice el hecho de que y es la *mínima* cota superior de E.

En consecuencia $y^n = x$, y esto completa la demostración.

Corolario *Si a y b son números reales positivos y n es un entero positivo, entonces*

$$(ab)^{1/n} = a^{1/n}b^{1/n}.$$

Demostración Haciendo $\alpha = a^{1/n}$, $\beta = b^{1/n}$. Se tiene

$$ab = \alpha^n\beta^n = (\alpha\beta)^n,$$

debido a que la multiplicación es conmutativa. [Axioma (M2) de la Definición 1.12]. Por lo tanto, la afirmación de unicidad del Teorema 1.21 muestra que

$$(ab)^{1/n} = \alpha\beta = a^{1/n}b^{1/n}.$$

1.22 Decimales Terminaremos esta parte, señalando la relación que existe entre los números reales y los decimales.

Sea $x > 0$ un número real, y n_0 el mayor entero, tal que $n_0 \leq x$.(*). Elegidos $n_0, n_1, \ldots, n_{k-1}$, sea n_k el mayor entero, para el cual

$$n_0 + \frac{n_1}{10} + \cdots + \frac{n_k}{10^k} \leq x.$$

Sea, además, E el conjunto formado por los números

(5) $$n_0 + \frac{n_1}{10} + \cdots + \frac{n_k}{10^k} \qquad (k = 0, 1, 2, \ldots).$$

En estas condiciones, x es el sup de E. El desarrollo decimal de x es

(6) $$n_0 \cdot n_1 n_2 n_3 \cdots.$$

Inversamente, para todo decimal con desarrollo infinito (6) el conjunto E de los números (5) está acotado superiormente, y (6) es el desarrollo decimal del sup de E.

Como nunca utilizaremos números decimales, no entraremos en un estudio detallado.

EL SISTEMA EXTENDIDO DE LOS NÚMEROS REALES

1.23 Definición El sistema extendido de los números reales está constituido por el campo real R al que se han añadido dos símbolos, $+\infty$ y $-\infty$.

(*)(Nótese que la existencia de n_0 depende de la propiedad Arquimediana de R.)

Se conservará el orden original en R, y se definirá

$$-\infty < x < +\infty$$

para cada $x \in R$.

Es evidente entonces, que $+\infty$ es una cota superior de cada subconjunto del sistema extendido de los números reales, y que cada subconjunto que no es vacío tiene una mínima cota superior. Por ejemplo, si E es conjunto que no es vacío ni acotado superiormente en R, entonces $E = +\infty$ en el sistema extendido de los números reales.

Las mismas observaciones se aplican exactamente a las cotas inferiores.

El sistema extendido de los números reales no forma un campo, pero por conveniencia se acostumbra lo siguiente:

(*a*) Si x es un número real, se verifica que $-\infty < x < +\infty$ y

$$x + \infty = +\infty, \qquad x - \infty = -\infty, \qquad \frac{x}{+\infty} = \frac{x}{-\infty} = 0.$$

(*b*) Si $x > 0$, será $x \cdot (+\infty) = +\infty$, $x \cdot (-\infty) = -\infty$.

(*c*) Si $x < 0$, es $x \cdot (+\infty) = -\infty$, $x \cdot (-\infty) = +\infty$.

Cuando es conveniente hacer distinción explícita entre los números reales por un lado y los símbolos $+\infty$ y $-\infty$ por otro, los primeros se llaman *finitos*.

EL CAMPO COMPLEJO

1.24 Definición Un *número complejo* es un par ordenado de números reales (a, b). "Ordenado" significa que (a, b) y (b, a) se consideran distintos si $a \neq b$.

Sean $x = (a, b)$, $y = (c, d)$ dos números complejos. Se escribe $x = y$ si y solamente si $a = c$ y $b = d$. (Nótese que esta definición no es por completo superflua; debe pensarse en la igualdad de los números racionales representados como cocientes de enteros.) Se define

$$x + y = (a + c, b + d),$$
$$xy = (ac - bd, ad + bc).$$

1.25 Teorema *Las definiciones anteriores para la adición y la multiplicación vuelven al conjunto de todos los números complejos un campo, con* (0, 0) *y* (1,0) *en lugar de* 0 *y* 1.

Demostración Simplemente se verificarán los axiomas de campo de la Definición 1.12. (Se usa la estructura de campo de R, por supuesto.)

Sean $x = (a, b)$, $y = (c, d)$, $z = (e, f)$.

(A1) es evidente.

(A2) $x + y = (a + c, b + d) = (c + a, d + b) = y + x$.

(A3)
$$\begin{aligned}(x + y) + z &= (a + c, b + d) + (e, f)\\ &= (a + c + e, b + d + f)\\ &= (a, b) + (c + e, d + f) = x + (y + z).\end{aligned}$$

(A4) $x + 0 = (a, b) + (0, 0) = (a, b) = x$.

(A5) Haciendo $-x = (-a, -b)$. Entonces $x + (-x) = (0, 0) = 0$.

(M1) es evidente.

(M2) $xy = (ac - bd, ad + bc) = (ca - db, da + cb) = yx$.

(M3)
$$\begin{aligned}(xy)z &= (ac - bd, ad + bc)(e, f)\\ &= (ace - bde - adf - bcf, acf - bdf + ade + bce)\\ &= (a, b)(ce - df, cf + de) = x(yz).\end{aligned}$$

(M4) $1x = (1, 0)(a, b) = (a, b) = x$.

(M5) Si $x \neq 0$, entonces $(a, b) \neq (0, 0)$, lo cual significa que al menos uno de los números reales a, b es diferente de 0. En consecuencia, por la Proposición 1.18(d) $a^2 + b^2 > 0$, y se puede definir

$$\frac{1}{x} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right).$$

Por tanto,

$$x \cdot \frac{1}{x} = (a, b)\left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right) = (1, 0) = 1.$$

(D)
$$\begin{aligned}x(y + z) &= (a, b)(c + e, d + f)\\ &= (ac + ae - bd - bf, ad + af + bc + be)\\ &= (ac - bd, ad + bc) + (ae - bf, af + be)\\ &= xy + xz.\end{aligned}$$

1.26 Teorema *Para números reales cualesquiera a y b se tiene*

$$(a, 0) + (b, 0) = (a + b, 0), \qquad (a, 0)(b, 0) = (ab, 0).$$

La demostración es trivial.

El Teorema 1.26 muestra que los números complejos de la forma $(a, 0)$ tienen las mismas propiedades aritméticas que los números reales correspon-

dientes a. Por tanto, es posible identificar $(a, 0)$ con a. Esta identificación hace del campo real un subcampo del campo complejo.

El lector puede haber notado que hasta ahora se han definido los números complejos sin hacer ninguna referencia a la misteriosa raíz cuadrada de -1. A continuación se muestra que la notación (a, b) es equivalente a la más acostumbrada $a + bi$.

1.27 Definición $i = (0, 1)$.

1.28 Teorema $i^2 = -1$.

Demostración $i^2 = (0, 1)(0, 1) = (-1, 0) = -1$.

1.29 Teorema *Si a y b son dos números reales, será $(a,b) = a + bi$.*

Demostración

$$\begin{aligned} a + bi &= (a, 0) + (b, 0)(0, 1) \\ &= (a, 0) + (0, b) = (a, b). \end{aligned}$$

1.30 Definición Si a, b son reales y $z = a + bi$, entonces al número complejo $\bar{z} = a - bi$ se le llama el *conjugado* de z. Los números a y b son la *parte real* y la *parte imaginaria* de z, respectivamente.

Se escribirá algunas veces

$$a = \text{Re}(z), \qquad b = \text{Im}(z).$$

1.31 Teorema *Si z y w son complejos, entonces*

(a) $\overline{z + w} = \bar{z} + \bar{w}$,
(b) $\overline{zw} = \bar{z} \cdot \bar{w}$,
(c) $z + \bar{z} = 2\,\text{Re}(z), z - \bar{z} = 2i\,\text{Im}(z)$,
(d) *$z\bar{z}$ es real y positivo (excepto cuando $z = 0$).*

Demostración (a), (b) y (c) son triviales. Para probar (d), escríbase $z = a + bi$, y nótese que $z\bar{z} = a^2 + b^2$.

1.32 Definición Si z es un número complejo, su *valor absoluto (o módulo)* $|z|$ es la raíz cuadrada no negativa de $z\bar{z}$; es decir, $|z| = (z\bar{z})^{1/2}$.

La existencia (y la unicidad) de $|z|$ se concluye a partir del Teorema 1.21 y la parte (d) del Teorema 1.31.

Nótese que cuando x es real, es $\bar{x} = x$, y por consiguiente $|x| = \sqrt{x^2}$. Así que $|x| = x$ si $x \geq 0$, $|x| = -x$ si $x < 0$.

1.33 Teorema *Siendo z y w números complejos. Se tiene*

(*a*) $|z| > 0$ a menos que $z = 0$, $|0| = 0$,
(*b*) $|\bar{z}| = |z|$,
(*c*) $|zw| = |z|\,|w|$,
(*d*) $|\text{Re}\, z| \leq |z|$,
(*e*) $|z + w| \leq |z| + |w|$.

Demostración (*a*) y (*b*) son evidentes. Si se hace $z = a + bi$, $w = c + di$, en donde a, b, c, d real. Entonces

$$|zw|^2 = (ac - bd)^2 + (ad + bc)^2 = (a^2 + b^2)(c^2 + d^2) = |z|^2|w|^2$$

o $|zw|^2 = (|z|\,|w|)^2$. Ahora (*c*) deduce de la afirmación de unicidad del Teorema 1.21.

Para demostrar (*d*), nótese que $a^2 \leq a^2 + b^2$, por consiguiente

$$|a| = \sqrt{a^2} \leq \sqrt{a^2 + b^2}.$$

Para demostrar (*e*), nótese que $\bar{z}w$ es el conjugado de $z\bar{w}$, así que $z\bar{w} + \bar{z}w = 2\, Re\,(z\bar{w})$. Por lo tanto

$$\begin{aligned} |z + w|^2 &= (z + w)(\bar{z} + \bar{w}) = z\bar{z} + z\bar{w} + \bar{z}w + w\bar{w} \\ &= |z|^2 + 2\,\text{Re}\,(z\bar{w}) + |w|^2 \\ &\leq |z|^2 + 2|z\bar{w}| + |w|^2 \\ &= |z|^2 + 2|z|\,|w| + |w|^2 = (|z| + |w|)^2. \end{aligned}$$

Y por último se obtiene (*e*) extrayendo raíz cuadrada en ambos miembros.

1.34 Notación Si $x_1, \ldots, x_n$ son números complejos, escribimos

$$x_1 + x_2 + \cdots + x_n = \sum_{j=1}^{n} x_j.$$

Terminamos esta sección con una desigualdad importante, corrientemente llamada desigualdad de Schwarz.

1.35 Teorema *Si $a_1, \ldots, a_n$ y $b_1, \ldots, b_n$ son números complejos, será*

$$\left|\sum_{j=1}^{n} a_j \bar{b}_j\right|^2 \leq \sum_{j=1}^{n} |a_j|^2 \sum_{j=1}^{n} |b_j|^2.$$

Demostración Pongamos $A = \Sigma|a_j|^2$; $B = \Sigma|b_j|^2$; $C = \Sigma a_j\bar{b}_j$ (en todas las sumas de esta demostración, j toma los valores $1,\dots n$). Si $B = 0$, será $b_1 = \cdots = b_n = 0$ y la conclusión es obvia. Supongamos, por tanto, que $B > 0$. Por el Teorema 1.31, tenemos

$$\begin{aligned}\sum |Ba_j - Cb_j|^2 &= \sum (Ba_j - Cb_j)(B\bar{a}_j - \overline{Cb_j})\\ &= B^2 \sum |a_j|^2 - B\bar{C} \sum a_j\bar{b}_j - BC\sum \bar{a}_j b_j + |C|^2 \sum |b_j|^2\\ &= B^2A - B|C|^2\\ &= B(AB - |C|^2).\end{aligned}$$

Como cada término de la primera suma es no negativo, vemos que

$$B(AB - |C|^2) \geq 0.$$

Como $B > 0$, se sigue que $AB - |C|^2 \geq 0$, que es la desigualdad deseada.

ESPACIOS EUCLIDIANOS

1.36 Definiciones Para cada entero positivo k, sea R^k el conjunto de todas las k-adas ordenadas

$$\mathbf{x} = (x_1, x_2, \dots, x_k),$$

donde $x_1, \dots, x_k$ son números reales, llamados *coordenadas* de $\mathbf{x}$. Los elementos de R^k se llaman puntos, o vectores, especialmente cuando $k > 1$. Designaremos a los vectores por letras negritas. Si $\mathbf{y} = (y_1, \dots, y_k)$ y α un número real, haremos

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_k + y_k),$$
$$\alpha\mathbf{x} = (\alpha x_1, \dots, \alpha x_k)$$

de modo que $\mathbf{x} + \mathbf{y} \in R^k$ y $\alpha\mathbf{x} \in R^k$. Esto define la adición de vectores, lo mismo que la multiplicación de un vector por un número real (un escalar). Estas dos operaciones satisfacen las leyes conmutativa, asociativa y distributiva (la demostración es elemental, comparando con las leyes análogas para los números reales) y hacen de R^k un *espacio vectorial sobre el campo real.* El elemento cero de R^k (a veces llamado *origen* o *vector nulo*) es el punto $\mathbf{0}$, cuyas coordenadas son todas 0.

Definimos también el llamado «producto interior» (o escalar) de $\mathbf{x}$ y $\mathbf{y}$ por

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{k} x_i y_i$$

y la *norma* de $\mathbf{x}$ por

$$|\mathbf{x}| = (\mathbf{x} \cdot \mathbf{x})^{1/2} = \left(\sum_{1}^{k} x_i^2\right)^{1/2}.$$

La estructura así definida (el espacio vectorial R^k con el producto interior y la norma) es el llamado k-espacio euclidiano.

1.37 Teorema *Supongamos* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in R^k$ *y que a es un número real. Será:*

(*a*) $|\mathbf{x}| \geq 0$;
(*b*) $|\mathbf{x}| = 0$, *si, y solo si,* $\mathbf{x} = \mathbf{0}$;
(*c*) $|\alpha\mathbf{x}| = |\alpha|\,|\mathbf{x}|$;
(*d*) $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}|\,|\mathbf{y}|$;
(*e*) $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$;
(*f*) $|\mathbf{x} - \mathbf{z}| \leq |\mathbf{x} - \mathbf{y}| + |\mathbf{y} - \mathbf{z}|$.

Demostración (*a*), (*b*) y (*c*) son evidentes y (*d*) es una consecuencia inmediata de la desigualdad de Schwarz. Por (*d*) tenemos

$$\begin{aligned} |\mathbf{x} + \mathbf{y}|^2 &= (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) \\ &= \mathbf{x} \cdot \mathbf{x} + 2\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y} \\ &\leq |\mathbf{x}|^2 + 2|\mathbf{x}|\,|\mathbf{y}| + |\mathbf{y}|^2 \\ &= (|\mathbf{x}| + |\mathbf{y}|)^2, \end{aligned}$$

con lo que queda demostrado (*e*). Finalmente, (*f*) se sigue de (*e*) si sustituimos $\mathbf{x}$ por $\mathbf{x} - \mathbf{y}$ y $\mathbf{y}$ por $\mathbf{y} - \mathbf{z}$.

1.38 Nota El Teorema 1.37(*a*), (*b*) y (*f*) nos permite (véase Cap. 2) considerar R^k como un espacio métrico.

R^1 (el conjunto de todos los números reales) suele llamarse recta, o recta real. Del mismo modo a R^2 se le llama plano, o plano complejo (comparar las Definiciones 1.24 y 1.36). En estos dos casos la norma coincide con el valor absoluto del número real o complejo correspondiente.

APÉNDICE

Se demostrará en este Apéndice el Teorema 1.19 construyendo R a partir de Q. Esta construcción se dividirá en varios pasos.

Paso 1 Los miembros de R serán determinados subconjuntos de Q, llamados *cortaduras*. Por definición, una cortadura es cualquier conjunto $\alpha \subset Q$ con las tres propiedades siguientes:

(I) α no es vacío, $\alpha \neq Q$.
(II) Si $p \in \alpha$, $q \in Q$, y $q < p$, entonces $q \in \alpha$.
(III) Si $p \in \alpha$, entonces $p < r$ para algún $r \in \alpha$.

Las letras $p, q, r, \ldots$ siempre denotarán números racionales, y α, β, $\gamma, \ldots$ se usarán para simbolizar cortaduras.

Nótese que (III) significa simplemente que α no tiene un miembro más grande; (II) implica dos hechos que se usarán con libertad:

Si $p \in \alpha$ y $q \notin \alpha$, entonces $p < q$.
Si $r \notin \alpha$ y $r < s$, entonces $s \notin \alpha$.

Paso 2 Al definir "$\alpha < \beta$" esto significará que: α es un subconjunto propio de β.

Comprobemos que esto cumple con los requisitos de la Definición 1.5.

Si $\alpha < \beta$ y $\beta < \gamma$, es evidente que $\alpha < \gamma$. (Un subconjunto propio de un subconjunto propio es un subconjunto propio.) También es claro que a lo más una de las tres relaciones siguientes

$$\alpha < \beta, \qquad \alpha = \beta, \qquad \beta < \alpha$$

puede ser cierta para cualquier par α, β. Para mostrar que al menos una es cierta, supóngase que las primeras dos no lo son. Entonces α no es un subconjunto de β. De aquí que hay un $p \in \alpha$ con $p \notin \beta$. Si $q \in \beta$, se deduce que $q < p$ (ya que $p \notin \beta$), en consecuencia $q \in \alpha$, por (II). Entonces $\beta \subset \alpha$. Debido a que $\beta \neq \alpha$, se concluye: $\beta < \alpha$.

Es por esto que ahora R es un conjunto ordenado.

Paso 3 *El conjunto ordenado R tiene la propiedad de la mínima cota superior.*

Para demostrar esto, supongamos que A es un subconjunto de R, y que $\beta \in R$ es una cota superior de A. Defínase a γ como la unión de todas las $\alpha \in A$. Dicho de otra forma, $p \in \gamma$ si y solamente si $p \in \alpha$ para alguna $\alpha \in A$. Se demostrará a continuación que $\gamma \in R$ y $\gamma = \sup A$.

Como A es no vacío, existe una $\alpha_0 \in A$. Esta α_0 no es vacía. Ya que $\alpha_0 \subset \gamma$, γ es no vacía. En seguida, $\gamma \subset \beta$ (porque $\alpha \subset \beta$ para cada $\alpha \in A$), y por lo tanto $\gamma \neq Q$. De aquí que γ satisface la propiedad (I). Para probar (II) y (III), tómese $p \in \gamma$. Entonces $p \in \alpha_1$ para alguna $\alpha_1 \in A$. Si $q < p$, entonces $q \in \alpha_1$, por esto $q \in \gamma$; esto demuestra (II). Si se selecciona $r \in \alpha_1$ de tal forma que $r > p$, se ve que $r \in \gamma$ (porque $\alpha_1 \subset \gamma$), y por lo tanto γ satisface (III).

Así que $\gamma \in R$.

Es evidente que $\alpha \leq \gamma$ para cada $\alpha \in A$.

Supóngase ahora que $\delta < \gamma$. Entonces hay un $s \in \gamma$ y $s \notin \delta$. Debido a que $s \in \gamma$, $s \in \alpha$ para alguna $\alpha \in A$. Por esto $\delta < \alpha$, y δ no es una cota superior de A.

Esto da el resultado deseado: $\gamma = \sup A$.

Paso 4 Si $\alpha \in R$ y $\beta \in R$, se define $\alpha + \beta$ como el conjunto de todas las sumas $r + s$, en donde $r \in \alpha$ y $s \in \beta$.

Se define también 0^* como el conjunto de todos los números racionales negativos. Es claro que 0^* es una cortadura. *Se verificará que los axiomas para la adición* (véase Definición 1.12) *son ciertos en R*, en donde 0^* desempeña el papel de 0.

(A1) Se tiene que mostrar que $\alpha + \beta$ es una cortadura. Es evidente que $\alpha + \beta$ es un subconjunto de Q que no es vacío. Si se toman $r' \notin \alpha$, $s' \notin \beta$. Entonces $r' + s' > r + s$ para cualesquiera $r \in \alpha$, $s \in \beta$. Es por esto que $r' + s' \notin \alpha + \beta$. De aquí que $\alpha + \beta$ tiene la propiedad (I).

Si se toma $p \in \alpha + \beta$. Entonces $p = r + s$, con $r \in \alpha$, $s \in \beta$. Si $q < p$, es $q - s < r$, así $q - s \in \alpha$, y $q = (q - s) + s \in \alpha + \beta$. Entonces (II) se cumple. Si se escoge $t \in \alpha$ de tal manera que $t > r$. Entonces $p < t + s$ y $t + s \in \alpha + \beta$. De donde (III) se cumple.

(A2) $\alpha + \beta$ es el conjunto de todos los $r + s$, con $r \in \alpha$, $s \in \beta$. Por la misma definición, $\beta + \alpha$ es el conjunto de todos los $s + r$. Debido a que $r + s = s + r$ para todo $r \in Q$, $s \in Q$, se tiene $\alpha + \beta = \beta + \alpha$.

(A3) Como en el punto anterior, esto se deriva de la ley asociativa en Q.

(A4) Si $r \in \alpha$ y $s \in 0^*$, entonces $r + s < r$, en consecuencia $r + s \in \alpha$. De aquí que $\alpha + 0^* \subset \alpha$. Para obtener la inclusión opuesta, tómese $p \in \alpha$, y $r \in \alpha$, $r > p$. Por lo que $p - r \in 0^*$, y $p = r + (p - r) \in \alpha + 0^*$. De donde $\alpha \subset \alpha + 0^*$. Se concluye que $\alpha + 0^* = \alpha$.

(A5) Con $\alpha \in R$ fijo y β el conjunto de todos los p con la propiedad siguiente:

Existe $r > 0$ tal que $-p - r \notin \alpha$.

Dicho de otra manera, algún número racional más pequeño que $-p$ ya no está en α.

Se mostrará que $\beta \in R$ y $\alpha + \beta = 0^*$.

Si $s \notin \alpha$ y $p = -s - 1$, entonces $-p - 1 \notin \alpha$, de aquí que $p \in \beta$. De esta manera β no es vacío. Si $q \in \alpha$, entonces $-q \notin \beta$. Así $\beta \neq Q$. En consecuencia, β satisface (I).

Tomando un $p \in \beta$, y $r > 0$, de manera que $-p - r \notin \alpha$. Si $q < p$, entonces $-q - r > -p - r$, de aquí que $-q - r \notin \alpha$. Por lo que $q \in \beta$, y (II) se cumple. Poniendo $t = p + (r/2)$. Entonces $t > p$, y $-t - (r/2) = -p - r \notin \alpha$, de manera que $t \in \beta$. En consecuencia β satisface (III).

Se ha probado que $\beta \in R$.

Si $r \in \alpha$ y $s \in \beta$, entonces $-s \notin \alpha$, por esto $r < -s$, $r + s < 0$. De aquí que $\alpha + \beta \subset 0^*$.

Para probar la inclusión opuesta, tómese $v \in 0^*$, y hágase $w = -v/2$. Entonces $w > 0$, y hay un entero n tal que $nw \in \alpha$, pero $(n+1)w \notin \alpha$. (Nótese que esto depende del hecho ¡de que Q tiene la propiedad arquimediana!) Poniendo $p = -(n+2)w$. Entonces $p \in \beta$, debido a que $-p - w \notin \alpha$, y

$$v = nw + p \in \alpha + \beta.$$

De aquí que $0^* \subset \alpha + \beta$.

Se concluye que $\alpha + \beta = 0^*$.

Por supuesto que esta β se denotará mediante $-\alpha$.

Paso 5 Habiendo demostrado que la adición definida en el Paso 4 satisface los axiomas (A) de la Definición 1.12, se concluye que la Proposición 1.14 es válida en R, y se puede probar uno de los requisitos de la Definición 1.17:

Si $\alpha, \beta, \gamma \in R$ *y* $\beta < \gamma$, *entonces* $\alpha + \beta < \alpha + \gamma$.

En verdad, es obvio de la definición de $+$ en R que $\alpha + \beta \subset \alpha + \gamma$; si se tuviera $\alpha + \beta = \alpha + \gamma$, la ley de la cancelación (Proposición 1.14) implicaría $\beta = \gamma$.

También se deduce que $\alpha > 0^*$ si y solamente si $-\alpha < 0^*$.

Paso 6 La multiplicación en este contexto es un poco más fastidiosa que la adición, debido a que los productos de racionales negativos son positivos. Por esta razón nos restringiremos primero a R^+, el conjunto de todas las $\alpha \in R$ con $\alpha > 0^*$.

Si $\alpha \in R^+$ y $\beta \in R^+$, entonces se define $\alpha\beta$ como el conjunto de todos los p tales que $p \leq rs$ para $r \in \alpha$, $s \in \beta$, $r > 0$, $s > 0$.

Se define 1^* como el conjunto de todos los $q < 1$.

Entonces se cumplen los axiomas (*M*) *y* (*D*) *de la Definición* 1.12, *con* R^+ *en lugar de* F *y* 1^* *desempeña el papel de* 1.

Las demostraciones son tan similares a las ofrecidas en el Paso 4 que se omitirán.

En particular, nótese que el segundo requisito de la Definición 1.17 se cumple: Si $\alpha > 0^*$ y $\beta > 0^*$, entonces $\alpha\beta > 0^*$.

Paso 7 Se completa la definición de la multiplicación estableciendo que $\alpha 0^* = 0^*\alpha = 0^*$, y

$$\alpha\beta = \begin{cases} (-\alpha)(-\beta) & \text{si } \alpha < 0^*, \beta < 0^*, \\ -[(-\alpha)\beta] & \text{si } \alpha < 0^*, \beta > 0^*, \\ -[\alpha \cdot (-\beta)] & \text{si } \alpha > 0^*, \beta < 0^*. \end{cases}$$

Los productos del miembro derecho se definieron en el Paso 6.

Habiendo probado (en el Paso 6) que los axiomas (M) se cumplen en R^+, ahora es muy simple demostrarlos en R, por medio de la aplicación repetida de la identidad $\gamma = -(-\gamma)$ que es parte de la Proposición 1.14. (Véase el Paso 5.)

La demostración de la ley distributiva

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$$

se divide en varias partes. Por ejemplo, supóngase que $\alpha > 0^*$, $\beta < 0^*$, $\beta + \gamma > 0^*$. Entonces $\gamma = (\beta + \gamma) + (-\beta)$, y (debido a que ya se sabe que la ley distributiva se cumple en R^+)

$$\alpha\gamma = \alpha(\beta + \gamma) + \alpha \cdot (-\beta).$$

Pero $\alpha \cdot (-\beta) = -(\alpha\beta)$. Entonces

$$\alpha\beta + \alpha\gamma = \alpha(\beta + \gamma).$$

Las otras partes se tratan en la misma forma.

Se ha completado ya la demostración de que R es un campo ordenado con la propiedad de la mínima cota superior.

Paso 8 Con cada $r \in Q$ se asocia el conjunto r^* que consta de todos los $p \in Q$ tales que $p < r$. Es evidente que cada r^* es una cortadura; es decir, $r^* \in R$. Estas cortaduras satisfacen las relaciones siguientes:

(*a*) $r^* + s^* = (r + s)^*$,
(*b*) $r^*s^* = (rs)^*$,
(*c*) $r^* < s^*$ si y solo si $r < s$.

Para probar (*a*) se escoge $p \in r^* + s^*$. Entonces $p = u + v$, en donde $u < r$, $v < s$. En consecuencia $p < r + s$, lo cual expresa que $p \in (r + s)^*$.

De manera inversa, supóngase que $p \in (r + s)^*$. Entonces $p < r + s$. Si se escoge t tal que $2t = r + s - p$, hágase

$$r' = r - t, s' = s - t.$$

Entonces $r' \in r^*$, $s' \in s^*$ y $p = r' + s'$, de modo que $p \in r^* + s^*$.

Esto demuestra (*a*). La demostración de (*b*) es similar.

Si $r < s$, entonces $r \in s^*$, pero $r \notin r^*$; en consecuencia $r^* < s^*$.

Si $r^* < s^*$ entonces hay un $p \in s^*$ tal que $p \notin r^*$. En consecuencia $r \leq p < s$, así que $r < s$.

Esto demuestra (*c*).

Paso 9 Como se vio en el Paso 8, el reemplazo de los números racionales r por las correspondientes "cortaduras racionales" $r^* \in R$ preserva las sumas, productos y el orden. Esto puede expresarse también, diciendo que el campo ordenado Q es *isomorfo* al campo ordenado Q^* cuyos elementos son las cortaduras racionales. Por supuesto que r^* no es de ninguna manera el mismo que r, pero las propiedades que nos interesan (la aritmética y el orden) son las mismas en los dos campos.

Esta identificación de Q con Q^* permite considerar a Q como un subcampo de R.

En términos de esta identificación, la segunda parte del Teorema 1.19 se entiende bien. Nótese que ocurre lo mismo cuando los números reales se consideran como subcampo del campo complejo, y esto sucede también en un nivel mucho más elemental, cuando los enteros se identifican como un subconjunto determinado de Q.

Es un hecho que *dos campos ordenados cualesquiera con la propiedad de la mínima cota superior son isomorfos*, pero esto no se demostrará aquí. Por lo tanto, la primera parte del Teorema 1.19 caracteriza completamente al campo real R.

Los libros de Landau y Thurston que se citan en la Bibliografía se dedican completamente a sistemas de números reales. El capítulo 1 del libro de Knopp contiene una descripción más pausada de cómo se puede obtener R a partir de Q. Otra construcción presentada en la sección 5 del libro de Hewitt y Stromberg, define a cada número real como una clase de equivalencia de sucesiones de Cauchy de números racionales (véase el Cap. 3).

Dedekind inventó las cortaduras que se usaron en esta sección. La construcción de R a partir de Q por medio de sucesiones de Cauchy se debe a Cantor. Cantor y Dedekind publicaron sus construcciones en 1872.

EJERCICIOS

A menos que se especifique lo contrario, todos los números que se mencionen en estos ejercicios serán reales.

1. Si r es racional ($r \neq 0$) y x es irracional, demuestre que $r + x$ y rx son irracionales.
2. Demostrar que no hay ningún número racional cuyo cuadrado sea 12.
3. Demostrar la Proposición 1.15.
4. Sea E un subconjunto que no es vacío de un conjunto ordenado; supóngase que α es una cota inferior de E y β es una cota superior de E. Demostrar que $\alpha \leq \beta$.
5. Sea A un conjunto de números reales que no es vacío, que está acotado inferiormente. Si $-A$ es el conjunto de todos los números $-x$, en donde $x \in A$. Demostrar que

$$\inf A = -\sup(-A).$$

6. Sea $b > 1$ fijo.

(*a*) Si m, n, p, q son enteros, $n > 0$, $q > 0$, y $r = m/n = p/q$, demostrar que

$$(b^m)^{1/n} = (b^p)^{1/q}.$$

Por consiguiente tiene sentido definir $b^r = (b^m)^{1/n}$.

(*b*) Demostrar que $b^{r+s} = b^r b^s$ si r y s son racionales.

(*c*) Si x es real, y se define $B(x)$ como el conjunto de todos los números b^t, donde t es racional y $t \leq x$. Demostrar que

$$b^r = \sup B(r)$$

en la cual r es racional. En consecuencia tiene sentido definir

$$b^x = \sup B(x)$$

para cada real x.

(*d*) Demostrar que $b^{x+y} = b^x b^y$ para todos los reales x y y.

7. Con $b > 1$, $y > 0$ fijos, demostrar que hay un real único x tal que $b^x = y$, al completar el bosquejo siguiente. (A este x se le llama *logaritmo de y de base b.*)

(*a*) Para cualquier entero positivo n, $b^n - 1 \geq n(b-1)$.

(*b*) En consecuencia $b - 1 \geq n(b^{1/n} - 1)$.

(*c*) Si $t > 1$ y $n > (b-1)/(t-1)$, entonces $b^{1/n} < t$.

(*d*) Si w es tal que $b^w < y$, entonces $b^{w+(1/n)} < y$ para un n suficientemente grande; para ver esto aplíquese la parte (*c*) con $t = y \cdot b^{-w}$.

(*e*) Si $b^w > y$, entonces $b^{w-(1/n)} > y$ para un n suficientemente grande.

(*f*) Considerar A como el conjunto de todos los w tales que $b^w < y$, y mostrar que $x = \sup A$ satisface $b^x = y$.

(*g*) Demostrar que x es único.

8. Demostrar que en el campo complejo no puede definirse ningún orden para que éste se vuelva un campo ordenado. *Sugerencia*: -1 es un cuadrado.

9. Supóngase que $z = a + bi$, $w = c + di$, y defínase $z < w$ si $a < c$, y también si $a = c$, pero cuando $b < d$. Demostrar que esto convierte al conjunto de los números complejos en un conjunto ordenado. (Por razones obvias, a este tipo de relación de orden se le denomina *orden diccionario o lexicográfico.*) ¿Tiene este conjunto ordenado la propiedad de la mínima cota superior?

10. Supóngase que $z = a + bi$, $w = u + iv$, y

$$a = \left(\frac{|w| + u}{2}\right)^{1/2}, \qquad b = \left(\frac{|w| - u}{2}\right)^{1/2}.$$

Demostrar que $z^2 = w$ si $u \geq 0$ y que $(\bar{z})^2 = w$ si $v \leq 0$. Y conclúyase que cada número complejo (con una excepción) tiene dos raíces cuadradas complejas.

11. Si z es un número complejo, demostrar que existe un $r \geq 0$ y un número complejo w con $|w| = 1$ tal que $z = rw$. ¿z determina siempre de manera única a w y r?

12. Si $z_1, \ldots, z_n$ son complejos, demostrar que

$$|z_1 + z_2 + \cdots + z_n| \leq |z_1| + |z_2| + \cdots + |z_n|.$$

13. Si x, y son complejos, demostrar que

$$||x| - |y|| \leq |x - y|.$$

14. Si z es un número complejo tal que $|z| = 1$, esto es, tal que $z\bar{z} = 1$, calcular

$$|1 + z|^2 + |1 - z|^2.$$

15. ¿En qué condiciones se produce la igualdad en la desigualdad de Schwarz?

16. Supongamos $k \geq 3$; $\mathbf{x}, \mathbf{y} \in R^k$; $|\mathbf{x} - \mathbf{y}| = d > 0$, y $r > 0$. Demostrar:
(*a*) Si $2r > d$ hay una infinidad de $\mathbf{z} \in R^k$ tales que

$$|\mathbf{z} - \mathbf{x}| = |\mathbf{z} - \mathbf{y}| = r.$$

(*b*) Si $2r = d$, hay solamente uno de tales $\mathbf{z}$.
(*c*) Si $2r < d$, no hay tales $\mathbf{z}$.
¿Cómo se modificarían estas afirmaciones si k fuera 2 o 1?

17. Demostrar que

$$|\mathbf{x} + \mathbf{y}|^2 + |\mathbf{x} - \mathbf{y}|^2 = 2|\mathbf{x}|^2 + 2|\mathbf{y}|^2$$

Si $x \in R^k$ y $\mathbf{y} \in R^k$. Interpretarlo geométricamente, como una propiedad de los paralelogramos.

18. Si $k \geq 2$ y $\mathbf{x} \in R^k$, demostrar que existe $\mathbf{y} \in R^k$ tal que $\mathbf{y} \neq 0$ pero $\mathbf{x} \cdot \mathbf{y} = 0$. ¿Es también verdad si $k = 1$?

19. Suponer $\mathbf{a} \in R^k$ y $\mathbf{b} \in R^k$. Hallar $\mathbf{c} \in R^k$ y $r > 0$ tales que

$$|\mathbf{x} - \mathbf{a}| = 2|\mathbf{x} - \mathbf{b}|$$

si, y solo si $|\mathbf{x} - \mathbf{c}| = r$.
(*Solución*: $3\mathbf{c} = 4\mathbf{b} - \mathbf{a}$, $3r = 2|\mathbf{b} - \mathbf{a}|$.)

20. Refiriéndose al Apéndice, supóngase que la propiedad (II) se omite de la definición de cortadura, conservando las mismas definiciones de orden y la adición. Mostrar que el conjunto ordenado resultante tiene la propiedad de la mínima cota superior y que la adición satisface los axiomas (A1) al (A4) (¡con un elemento cero ligeramente diferente!) pero no el (A5).

2

TOPOLOGÍA BÁSICA

CONJUNTOS FINITOS, NUMERABLES Y NO NUMERABLES

Empezamos esta sección con una definición del concepto de función.

2.1 Definición Consideremos dos conjuntos A y B, cuyos elementos pueden ser objetos cualesquiera, y supongamos que con cada elemento x de A se asocia, de algún modo, un elemento de B que representaremos por $f(x)$. Se dice que f es una *función* de A en B (o una *aplicación* o *mapeo* de A en B).

El conjunto A se llama *dominio de definición* de f (también se dice que f está definida en A) y los elementos de $f(x)$ se llaman *valores* de f. El conjunto de todos los valores de f se llama *rango* de f.

2.2 Definición Sean A y B dos conjuntos y f un mapeo o aplicación de A en B. Si $E \subset A$, se define $f(E)$ como el conjunto de todos los elementos $f(x)$ para $x \in E$. A $f(E)$ le llamamos *imagen* de E bajo f. En esta notación, $f(A)$ es el rango de f. Está claro que $f(A) \subset B$. Si $f(A) = B$, decimos que f mapea o aplica A *sobre* B. (Se utiliza la palabra *sobre*, admitiendo para ella un significado más específico que el de *en*).

Si $E \subset B$, $f^{-1}(E)$ representa el conjunto de todo $x \in A$ tal que $f(x) \in E$. Llamamos a $f^{-1}(E)$ *imagen inversa* de E bajo f. Si $y \in B$, $f^{-1}(y)$ es el conjunto de todos los $x \in A$ tales que $f(x) = y$. Si, para cada $y \in B$, $f^{-1}(y)$ no

está integrado por más de un elemento de A se dice que f es un mapeo 1-1 (*uno a uno*) de A en B. Esto puede expresarse también como sigue: f es un mapeo 1-1 de A en B cuando $f(x_1) \neq f(x_2)$ siempre que $x_1 \neq x_2$; $x_1 \in A$; $x_2 \in A$.

(La notación $x_1 \neq x_2$ significa que x_1 y x_2 son elementos diferentes; en otro caso, escribiremos $x_1 = x_2$).

2.3 Definición Si existe un mapeo 1-1 de A *sobre* B, decimos que A y B pueden ponerse en *correspondencia* 1-1, también llamada *biunívoca*, que A y B tienen el mismo *número cardinal*, o más brevemente, que A y B son equivalentes, y escribimos $A \sim B$. Esta relación tiene las propiedades siguientes, como se ve claramente:

Es reflexiva: $A \sim A$.
Es simétrica: si $A \sim B$, también $B \sim A$.
Es transitiva: si $A \sim B$ y $B \sim C$, también $A \sim C$.

Toda relación con estas tres propiedades, se llama *relación de equivalencia*.

2.4 Definición Para todo entero positivo n, sea J_n el conjunto cuyos elementos son los números enteros $1, 2, \ldots, n$, y J el conjunto formado por todos los enteros positivos. Para todo conjunto A, decimos:

(*a*) A es *finito* si $A \sim J_n$ para algún n (el conjunto vacío se considera finito).
(*b*) A es *infinito* si no es finito.
(*c*) A es *numerable* si $A \sim J$.
(*d*) A es *no numerable* si no es ni finito ni numerable.
(*e*) A es *a lo más numerable* si es finito o numerable.

Para dos conjuntos finitos A y B, evidentemente tenemos $A \sim B$ si, y solo si, A y B contienen el mismo número de elementos. Para los conjuntos infinitos, la idea «tener el mismo número de elementos» es vaga, mientras que la noción de correspondencia 1-1 conserva su claridad.

2.5 Ejemplo Sea A el conjunto de todos los números enteros. A es numerable. Para verlo, consideremos la disposición siguiente de los conjuntos A y J:

$$A: \quad 0, 1, -1, 2, -2, 3, -3, \ldots$$
$$J: \quad 1, 2, 3, 4, 5, 6, 7, \ldots$$

Podemos, en este ejemplo, dar una fórmula explícita para una función f de J en A que establece una correspondencia 1-1:

$$f(n) = \begin{cases} \dfrac{n}{2} & (n \text{ par}), \\ -\dfrac{n-1}{2} & (n \text{ impar}). \end{cases}$$

2.6 Observación Un conjunto finito no puede ser equivalente a uno de sus subconjuntos propios. Sin embargo, esto es posible para conjuntos infinitos, según se ve en el ejercicio 2.5, en el que J es un subconjunto propio de A.

De hecho, podemos sustituir la Definición 2.4(b) por la proposición: A es infinito, si es equivalente a uno de sus subconjuntos propios.

2.7 Definición Por *sucesión* entendemos una función f definida en el conjunto J de todos los enteros positivos. Si $f(n) = x_n$, para $n \in J$, se acostumbra a representar la sucesión f por el símbolo $\{x_n\}$, o a veces por x_1, x_2, $x_3, \ldots$ Los valores de f, esto es, los elementos x_n, se llaman *términos* de la sucesión. Si A es un conjunto y $x_n \in A$ para todo $n \in J$, se dice que $\{x_n\}$ es una *sucesión en A*, o una *sucesión de elementos de A*.

Obsérvese que los términos x_1, x_2, $x_3, \ldots$ de una sucesión no necesitan ser distintos.

Como todo conjunto numerable es el rango de una función 1-1 definida en J, podemos considerar un conjunto numerable como el rango de una sucesión de términos distintos. Hablando más libremente, podemos decir que los elementos de un conjunto numerable pueden ser «dispuestos en una sucesión».

A veces es conveniente sustituir J en esta definición por el conjunto de todos los enteros no negativos, esto es, comenzar con 0 en lugar de 1.

2.8 Teorema *Todo subconjunto infinito de un conjunto numerable A, es numerable.*

Demostración Supongamos $E \subset A$ y que E es infinito. Dispongamos los elementos x de A en una sucesión $\{x_n\}$ de elementos distintos. Construyamos una sucesión $\{n_k\}$ como sigue:

Sea n_1, el menor entero positivo tal que $x_{n_1} \in E$. Elegidos $n_1, \ldots, n_{k-1}$ ($k = 2, 3, 4, \ldots$), sean n_k el menor entero mayor que n_{k-1} y tal que $x_{n_k} \in E$.

Poniendo $f(k) = x_{n_k}$ ($k = 1, 2, 3, \ldots$) obtenemos una correspondencia 1-1 entre E y J.

El teorema muestra que, hablando vulgarmente, los conjuntos numerables representan la «menor» infinidad; un conjunto que no es no numerable puede ser un subconjunto de uno numerable.

2.9 Definición Sean A y Ω dos conjuntos y supongamos que a cada elemento α de A hay asociado un subconjunto de Ω que representaremos por E_α.

El conjunto cuyos elementos son los conjuntos E_α se representará por $\{E_\alpha\}$. En lugar de hablar de conjuntos de conjuntos, hablaremos a veces de una colección de conjuntos, o de una familia de conjuntos.

Se define la *unión* de los conjuntos E_α como el conjunto S tal que $x \in S$ si, y solo si, $x \in E_\alpha$ para, al menos, un $\alpha \in A$. Usaremos la notación

$$S = \bigcup_{\alpha \in A} E_\alpha. \tag{1}$$

Si A está constituido por los enteros $1, 2, \dots, n$ escribiremos

$$S = \bigcup_{m=1}^{n} E_m \tag{2}$$

o

$$S = E_1 \cup E_2 \cup \cdots \cup E_n. \tag{3}$$

Si A es el conjunto de todos los enteros positivos, la notación usual es

$$S = \bigcup_{m=1}^{\infty} E_m. \tag{4}$$

El símbolo ∞ indica simplemente en (4), que se toma la unión de una colección *numerable* de conjuntos, y no debe confundirse con los símbolos $+\infty$ y $-\infty$ introducidos en la Definición 1.23.

Se define la *intersección* de los conjuntos E_α como el conjunto P tal que $x \in P$ si, y solo si, $x \in E_\alpha$ para todo $\alpha \in A$. Usaremos la notación

$$P = \bigcap_{\alpha \in A} E_\alpha, \tag{5}$$

o

$$P = \bigcap_{m=1}^{n} E_m = E_1 \cap E_2 \cap \cdots \cap E_n, \tag{6}$$

o

$$P = \bigcap_{m=1}^{\infty} E_m, \tag{7}$$

como para las uniones. Si $A \cap B$ no es vacío, decimos que A y B se *intersecan*; o de otro modo: son *ajenos*.

2.10 Ejemplos

(a) Supongamos que E_1 está constituido por 1, 2, 3, y E_2 por 2, 3, 4.

$E_1 \cup E_2$ estará constituido por 1, 2, 3, 4; mientras que $E_1 \cap E_2$ lo estará por 2, 3.
(*b*) Sea A el conjunto de los números reales x tales que $0 < x \leq 1$. Para todo $x \in A$, sea E_x el conjunto de los números reales y tales que $0 < y < x$. Será

(i) $E_x \subset E_z$ si, y solo si, $0 < x \leq z \leq 1$;

(ii) $\bigcup_{x \in A} E_x = E_1$;

(iii) $\bigcap_{x \in A} E_x$ es vacío;

(i) y (ii) son inmediatas. Para demostrar (iii), notemos que para todo $y > 0$ $y \notin E_x$ si $x < y$. De aquí, que $y \notin \bigcap_{x \in A} E_x$.

2.11 Observación Muchas propiedades de las uniones e intersecciones son completamente similares a las de las sumas y productos; a veces incluso se utilizan las palabras suma y producto y se escriben los símbolos Σ y Π en lugar de $\bigcup$ y $\bigcap$.

Las leyes conmutativas y asociativa se ven inmediatamente:

(8) $$A \cup B = B \cup A; \qquad A \cap B = B \cap A.$$

(9) $$(A \cup B) \cup C = A \cup (B \cup C); \qquad (A \cap B) \cap C = A \cap (B \cap C).$$

Así, la omisión de paréntesis en (3) y (6) está justificada.

También se conserva la ley distributiva:

(10) $$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Para demostrar esto, representemos el primer y segundo miembro de (10) por E y F, respectivamente.

Supongamos $x \in E$. Entonces; $x \in A$ y $x \in B \cup C$, esto es $x \in B$ o $x \in C$ (es posible que en ambos). Por tanto, $x \in A \cap B$ o $x \in A \cap C$, de modo que $x \in F$. Así pues, $E \subset F$.

Continuando: supongamos que $x \in F$. Entonces, $x \in A \cap B$ o $x \in A \cap C$. Esto es, $x \in A$, y $x \in B \cup C$. Por tanto, $x \in A \cap (B \cup C)$, de modo que $F \subset E$.

De lo cual se deduce que $F = E$.

Reseñemos algunas otras relaciones que pueden demostrarse fácilmente:

(11) $$A \subset A \cup B,$$

(12) $$A \cap B \subset A.$$

Si 0 representa un conjunto vacío, será

(13) $$A \cup 0 = A, \qquad A \cap 0 = 0.$$

Si $A \subset B$:

(14) $$A \cup B = B, \qquad A \cap B = A.$$

2.12 Teorema *Sea $\{E_n\}$, $n = 1, 2, 3, \ldots$ una sucesión de conjuntos numerables y hagamos*

(15) $$S = \bigcup_{n=1}^{\infty} E_n.$$

S es numerable.

Demostración Dispongamos cada conjunto E_n en una sucesión $\{x_{nk}\}$, con $k = 1, 2, 3, \ldots$ y consideremos la disposición en cuadro infinito:

(16) $$\begin{array}{ccccc} x_{11} & x_{12} & x_{13} & x_{14} & \cdots \\ x_{21} & x_{22} & x_{23} & x_{24} & \cdots \\ x_{31} & x_{32} & x_{33} & x_{34} & \cdots \\ x_{41} & x_{42} & x_{43} & x_{44} & \cdots \end{array}$$

en la que los elementos de E_n constituyen la fila n-sima. El cuadro contiene todos los elementos de S. Estos elementos pueden disponerse en una sucesión como indican las flechas.

(17) $$x_{11}; x_{21}, x_{12}; x_{31}, x_{22}, x_{13}; x_{41}, x_{32}, x_{23}, x_{14}; \ldots$$

Si dos cualesquiera de estos conjuntos E_n tienen elementos comunes, aparecerán más de una vez en (17). Por tanto, hay un subconjunto T del conjunto de todos los enteros positivos, tal que $S \sim T$, lo que demuestra que S es a lo sumo numerable (Teorema 2.8). Como $E_1 \subset S$ y E_1 es infinito, S es infinito y, por tanto, numerable.

Corolario *Supongamos que A es a lo sumo numerable, y para cada $a \in A$, B_α es a lo sumo numerable. Hagamos*

$$T = \bigcup_{\alpha \in A} B_\alpha.$$

T será también a lo sumo numerable.

Puesto que T es equivalente a un subconjunto de (15).

2.13 Teorema *Sea A un conjunto numerable y B_n el conjunto de todas las n-tuplas $(a_1, \ldots, a_n)$, donde $a_k \in A$ ($k = 1, \ldots n$) sin que los elementos $a_1, \ldots, a_n$ necesiten ser distintos. B_n es numerable.*

Demostración Que B_1 es numerable es evidente, pues $B_1 = A$. Su-

pongamos que B_{n-1} es numerable ($n = 2, 3, 4, \ldots$). Los elementos de B_n son de la forma

$$(b, a) \qquad (b \in B_{n-1}, a \in A). \tag{18}$$

Para cada b dado, el conjunto de pares (b,a) es equivalente a A, y por tanto numerable. Así pues, B_n es la unión de un conjunto numerable de conjuntos numerables. Por el Teorema 2.12, B_n es numerable.

El teorema se demuestra por inducción.

Corolario *El conjunto de todos los números racionales es numerable.*

Demostración Aplicaremos el Teorema 2.13, con $n = 2$, haciendo observar que todo número racional r es de la forma b/a, siendo a y b enteros. El conjunto de pares (a,b) y por tanto el de las fracciones b/a es numerable.

De hecho, también es numerable el conjunto de todos los números algebraicos (ver Ejer. 2).

Que no todos los conjuntos infinitos son numerables, se ve en el siguiente teorema.

2.14 Teorema *Sea A el conjunto de todas las sucesiones cuyos elementos son los dígitos* 0 *y* 1. *Este conjunto A, es no numerable.*

Los elementos de A se disponen en sucesión como sigue: 1, 0, 0, 1, 0, 1, 1, 1, ...

Demostración Sea E un subconjunto numerable de A y supongamos E constituido por las sucesiones $s_1, s_2, s_3, \ldots$ Construimos una sucesión s como sigue: si el dígito n-ésimo en s_n es 1, hacemos que el n-ésimo de s sea 0, y viceversa. La sucesión s difiere, pues, de cada elemento de E al menos en un lugar; por tanto, $s \notin E$. Pero es claro que $s \in A$, por lo cual E es un subconjunto propio de A.

Hemos demostrado que todo subconjunto numerable de A es un subconjunto propio de A. Se deduce de aquí que A es no numerable (porque de otro modo A sería un subconjunto propio de A, lo que es absurdo).

La idea de la demostración anterior fue utilizada primeramente por Cantor y se llama proceso diagonal de Cantor, porque si las sucesiones $s_1, s_2, s_3, \ldots$ están colocadas en una disposición como en (16), los elementos de la diagonal son los que intervienen en la construcción de la nueva sucesión.

Los lectores que estén familiarizados con la representación binaria de los números reales (base 2 en lugar de 10) se darán cuenta de que el Teorema

2.14 implica que el conjunto de todos los números reales es no numerable. Daremos una segunda demostración de este hecho en el Teorema 2.43.

ESPACIOS MÉTRICOS

2.15 Definición Un conjunto X cuyos elementos llamaremos *puntos*, se dice que es un *espacio métrico* si a cada dos puntos p y q de X hay asociado un número real $d(p,q)$ llamado distancia de p a q, tal que

(a) $d(p,q) > 0$ si $p \neq q$; $d(p,p) = 0$;
(b) $d(p,q) = d(q,p)$;
(c) $d(p,q) \leq d(p,r) + d(r,q)$, para todo $r \in X$.

Cualquier función con las tres propiedades anteriores se llama *función distancia* o *métrica*.

2.16 Ejemplos Los ejemplos más importantes de espacios métricos, desde nuestro punto de vista, son los espacios euclidianos R^k, especialmente R^1 (la recta real) y R^2 (el plano complejo); la distancia en R^k se define por

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}| \qquad (\mathbf{x}, \mathbf{y} \in R^k). \tag{19}$$

Por el Teorema 1.37, las condiciones de la definición 2.15 quedan satisfechas por (19).

Es importante observar que todo subconjunto Y de un espacio métrico X, es a su vez un espacio métrico, con la misma función distancia; porque está claro que si se cumplen las condiciones (a) a (c) de la Definición 2.15 para p, q, $r \in X$, también se cumplirán si imponemos a p, q, r la condición restrictiva de pertenecer a Y.

Así pues, todo subconjunto de un espacio euclidiano, es un espacio métrico. Otros ejemplos son los espacios $\mathscr{C}(K)$ y $\mathscr{L}^2(\mu)$, que se tratan en los capítulos 7 y 11, respectivamente.

2.17 Definición Por *segmento* (a,b) queremos significar el conjunto de todos los números reales x tales que $a < x < b$.

Por *intervalo* $[a,b]$ entendemos el conjunto de todos los números reales x tales que $a \leq x \leq b$.

Ocasionalmente encontraremos «intervalos semi-abiertos» $[a,b)$ y $(a,b]$, el primero de los cuales está constituido por todo x tal que $a \leq x < b$, y el segundo por todo x para el cual $a < x \leq b$.

Si $a_i < b_i$ para $i = 1,\ldots, k$, el conjunto de todos los puntos, $\mathbf{x} = (x_1,\ldots, x_k)$ en R^k, cuyas coordenadas satisfacen las desigualdades $a_i \leq x_i \leq b_i$ $(1 \leq i \leq k)$ se llama celda-k. Así, una celda-1 es un intervalo, una celda-2 un rectángulo, etc.

Si $\mathbf{x} \in R^k$ y $r > 0$, se define la bola *abierta (o cerrada)* B con centro en $\mathbf{x}$ y radio r como el conjunto de todo $\mathbf{y} \in R^k$ tal que $|\mathbf{y} - \mathbf{x}| < r$ (o $|\mathbf{y} - \mathbf{x}| \leq r$).

Llamaremos *convexo* a un conjunto $E \subset R^k$, si

$$\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in E$$

cuando $\mathbf{x} \in E$, $\mathbf{y} \in E$ y $0 < \lambda < 1$.

Por ejemplo, *las bolas son convexas*. Porque si $|\mathbf{y} - \mathbf{x}| < r$ $|\mathbf{z} - \mathbf{x}| < r$, y $0 < \lambda < 1$, tenemos

$$\begin{aligned}|\lambda\mathbf{y} + (1 - \lambda)\mathbf{z} - \mathbf{x}| &= |\lambda(\mathbf{y} - \mathbf{x}) + (1 - \lambda)(\mathbf{z} - \mathbf{x})| \\ &\leq \lambda|\mathbf{y} - \mathbf{x}| + (1 - \lambda)|\mathbf{z} - \mathbf{x}| < \lambda r + (1 - \lambda)r = r.\end{aligned}$$

La misma demostración se aplica a las bolas cerradas. Es también fácil de ver que las celdas-k son convexas.

2.18 Definición Sea X un espacio métrico. Se entiende que todos los puntos y conjuntos mencionados a continuación son elementos y subconjuntos de X.

(*a*) Vecindad (o *entorno*) de un punto p es un conjunto $N_r(p)$ formado por todos los puntos q tales que $d(p,q) < r$. Al número r se le llama *radio* de $N_r(p)$.

(*b*) Un punto p es un *punto límite* del conjunto E si toda vecindad de p contiene un punto $q \neq p$ tal que $q \in E$.

(*c*) Si $p \in E$ y p no es un punto límite de E, a p se le llama *punto aislado* de E.

(*d*) E es *cerrado* si todos sus puntos límites pertenecen a él.

(*e*) Un punto p es *interior* a E si existe una vecindad N de p tal que $N \subset E$.

(*f*) E es *abierto* si todos sus puntos son interiores.

(*g*) El *complemento de E* (representado por E^c) es el conjunto de todos los puntos $p \in X$ tales que $p \notin E$.

(*h*) E es *perfecto* si es cerrado y todos sus puntos son puntos límites.

(*i*) E es *acotado* si hay un número real M y un punto $q \in X$ tales que $d(p,q) < M$ para todo $p \in E$.

(*j*) E es *denso en X* si todo punto de X es punto límite de E, o punto de E (o ambas cosas a la vez).

Observemos que en R^1 las vecindades son segmentos, mientras que en R^2 son círculos.

2.19 Teorema *Toda vecindad es un conjunto abierto.*

Demostración Consideremos una vecindad $E = N_r(p)$, y sea q un punto cualquiera de E. Hay un número real positivo h, tal que

$$d(p, q) = r - h.$$

Para todo punto s para el cual $d(q,s) < h$ tendremos, pues

$$d(p, s) \leq d(p, q) + d(q, s) < r - h + h = r,$$

de modo que $s \in E$. Así pues, q es un punto interior de E.

2.20 Teorema *Si p es un punto límite de un conjunto E, toda vecindad de p contiene infinitos puntos de E.*

Demostración Supongamos que hay una vecindad N de p que solamente contiene un número finito de puntos de E. Sean $q_1, \ldots, q_n$ esos puntos de $N \cap E$ que son distintos de p, y hagamos

$$r = \min_{1 \leq m \leq n} d(p, q_m)$$

[usaremos esta notación para expresar el menor de los números $d(p,q_1), \ldots, d(p,q_n)$]. El mínimo de un conjunto finito de números positivos es, evidentemente, positivo, de modo que $r > 0$.

La vecindad $N_r(p)$ no contiene ningún punto q de E tal que $q \neq p$, de modo que p no es punto límite de E. Esta contradicción, demuestra el teorema.

Corolario *Un conjunto finito de puntos no tiene puntos límites.*

2.21 Ejemplos Consideremos los siguientes subconjuntos de R^2:

(*a*) El conjunto de todos los números complejos z tales que $|z| < 1$.
(*b*) El conjunto de todos los números complejos z tales que $|z| \leq 1$.
(*c*) Un conjunto finito.
(*d*) El conjunto de todos los enteros.
(*e*) El conjunto constituido por los números $1/n$ ($n = 1, 2, 3, \ldots$). Notemos que este conjunto E tiene un punto límite (éste es, $z = 0$), pero ningún punto de E es punto límite, con lo que se ve la diferencia entre poseer un punto límite y contenerle.
(*f*) El conjunto de todos los números complejos (esto es, R^2).
(*g*) El segmento (a,b).

Observemos que (*d*), (*e*) y (*g*) pueden ser considerados también como subconjuntos de R^1.

A continuación se expresan en una tabla algunas propiedades de estos conjuntos:

	Cerrado	*Abierto*	*Perfecto*	*Acotado*
(*a*)	No	Sí	No	Sí
(*b*)	Sí	No	Sí	Sí
(*c*)	Sí	No	No	Sí
(*d*)	Sí	No	No	No
(*e*)	No	No	No	Sí
(*f*)	Sí	Sí	Sí	No
(*g*)	No		No	Sí

En (*g*) hemos dejado la segunda columna en blanco. La razón es que el segmento (a,b) no es abierto si le consideramos como un subconjunto de R^2, pero si es un subconjunto abierto de R^1.

2.22 Teorema *Sea* $\{E_\alpha\}$ *una colección (finita o infinita) de conjuntos* E_α. *Será*:

(20)
$$\left(\bigcup_\alpha E_\alpha\right)^c = \bigcap_\alpha (E_\alpha^c).$$

Demostración Sean A y B el primero y el segundo miembro de (20). Si $x \in A$, entonces $x \notin \bigcup_\alpha E_\alpha$; por tanto $x \notin E_\alpha$ para cualquier α, y $x \in E_\alpha^c$ para todo α, de modo que $x \in \bigcap E_\alpha^c$. Así pues, $A \subset B$.

Inversamente, si $x \in B$, estará $x \in E_\alpha^c$ para todo α, y $x \notin E_\alpha$ para cualquier α, y de aquí $x \notin \bigcup_\alpha E_\alpha$, de modo que $x \in (\bigcup_\alpha E_\alpha)^c$. Así pues, $B \subset A$.

Se deduce pues, que $A = B$.

2.23 Teorema *Un conjunto E es abierto, si y solo si su complemento es cerrado.*

Demostración Supongamos, primeramente que E^c es cerrado, y elijamos un $x \in E$. En este caso, $x \notin E^c$ y x no es punto límite de E^c. Por tanto, existe una vecindad N de x tal que $E^c \cap N$ es vacío, esto es, $N \subset E$. Este x es un punto interior de E y E es abierto.

Ahora, supongamos E abierto y sea x un punto límite de E^c. Cada vecindad de x contiene un punto de E^c, de modo que x no es punto interior de E. Como E es abierto, esto significa que $x \in E^c$, de donde se deduce que E^c es cerrado.

Corolario *Un conjunto F es cerrado si, y solo si, su complemento es abierto.*

2.24 Teorema

(*a*) *Para toda colección* $\{G_\alpha\}$ *de conjuntos abiertos,* $\bigcup_\alpha G_\alpha$ *es abierto.*
(*b*) *Para toda colección* $\{F_\alpha\}$ *de conjuntos cerrados,* $\bigcap_\alpha F_\alpha$ *es cerrado.*
(*c*) *Para toda colección finita* $G_1, \ldots, G_n$ *de conjuntos abiertos,* $\bigcap_{i=1}^n G_i$ *es abierto.*
(*d*) *Para toda colección finita* $F_1, \ldots, F_n$ *de conjuntos cerrados,* $\bigcup_{i=1}^n F_i$ *es cerrado.*

Demostración Hagamos $G = \bigcup_\alpha G_\alpha$. Si $x \in G$, tendremos que $x \in G_\alpha$ para algún α. Como x es un punto interior de G_α, es también un punto interior de G, y G es abierto, lo que demuestra (*a*).

Por el Teorema 2.22,

(21) $$\left(\bigcap_\alpha F_\alpha\right)^c = \bigcup_\alpha (F_\alpha^c),$$

y F_α^c es abierto, por el Teorema 2.23. Por tanto (*a*) implica que (21) es abierto y, por tanto, $\bigcap_\alpha F_\alpha$ cerrado.

Ahora, hagamos $H = \bigcap_{i=1}^n G_i$. Para todo $x \in H$, existen vecindades N_i de x, con radios r_i, tales que $N_i \subset G_i$ $(i = 1, \ldots, n)$. Hagamos

$$r = \text{mín}\,(r_1, \ldots, r_n),$$

y sea N la vecindad de x de radio r. Entonces, $N \subset G_i$ para $i = 1, \ldots, n$ de forma que $N \subset H$, y H es abierto.

Tomando complementos de (*c*) se deduce (*d*):

$$\left(\bigcup_{i=1}^n F_i\right)^c = \bigcap_{i=1}^n (F_i^c).$$

2.25 Ejemplos En los apartados (*c*) y (*d*) del teorema precedente es esencial el carácter finito de las colecciones. Para verlo, sea G_n el segmento $\left(-\frac{1}{n}, \frac{1}{n}\right)$ $(n = 1, 2, 3, \ldots)$: G_n es un subconjunto abierto de R^1. Hagamos $G = \bigcap_{n=1}^\infty G_n$; G estará constituido por un solo punto (esto es, $x = 0$) y no es, por consiguiente, un subconjunto abierto de R^1.

Así, pues, la intersección de una colección infinita de conjuntos abiertos no es necesariamente abierta. Del mismo modo, la unión de una colección infinita de conjuntos cerrados no es necesariamente cerrada.

2.26 Definición Si X es un espacio métrico, $E \subset X$ y E' representa al conjunto de todos los puntos límite de E en X, entonces la *cerradura** de E es el conjunto $\bar{E} = E \cup E'$.

*N. del E.: En España y Sudamérica, *clausura*.

2.27 Teorema *Si X es un espacio métrico y $E \subset X$, entonces*

(*a*) *$\bar{E}$ es cerrado,*
(*b*) *$E = \bar{E}$ si, y solo si, E es cerrado,*
(*c*) *$\bar{E} \subset F$ para cada conjunto cerrado $F \subset X$ tal que $E \subset F$.*

De (*a*) y (*c*) es fácil ver que $\bar{E}$ es el subconjunto cerrado más *pequeño* de X que contiene E.

Demostración
(*a*) Si $p \in X$ y $p \notin \bar{E}$, entonces p no es punto de E ni tampoco punto límite de E. Por consiguiente, p tiene una vecindad que no intersecta E. Por tanto el complemento de $\bar{E}$ es abierto. De aquí que $\bar{E}$ es cerrado.
(*b*) Si $E = \bar{E}$, (*a*) implica que E es cerrado. Si E es cerrado, entonces $E' \subset E$ [por las Definiciones 2.18(*d*) y 2.26]; por lo tanto $\bar{E} = E$.
(*c*) Si F es cerrado y $F \supset E$, entonces $F \supset F'$, de aquí que $F \supset E'$. Esto es $F \supset \bar{E}$.

2.28 Teorema *Sea E un conjunto de números reales acotado superiormente y que no es vacío. Si $y = \sup E$. Entonces $y \in \bar{E}$. Por consiguiente $y \in E$ si E es cerrado.*

Compárese esto con los ejemplos de la sección 1.9.

Demostración Si $y \in E$, entonces $y \in \bar{E}$. Supóngase que $y \notin E$. Entonces para cada $h > 0$ existe un punto $x \in E$ tal que $y - h < x < y$, porque de otra forma $y - h$ sería una cota superior de E. De aquí que y es un punto límite de E. Por consiguiente, $y \in \bar{E}$.

2.29 Observación Supongamos $E \subset Y \subset X$, siendo X un espacio métrico. Decir que E es un subconjunto abierto de X significa que a cada punto $p \in E$ hay asociado un número positivo r, tal que las condiciones $d(p,q) < r$ y $q \in X$ implica que $q \in E$. Pero hemos visto anteriormente (Sec. 2.16) que Y es también un espacio métrico, por lo que nuestras definiciones se pueden hacer igual con Y. Para ser más explícitos, diremos que E es *abierto relativo en Y* si a cada $p \in E$ hay asociado un $r > 0$, tal que $q \in E$ cuando $d(p,q) < r$ y $q \in Y$. El Ejemplo 2.21(*g*) mostró que un conjunto puede ser abierto relativo en Y sin ser un subconjunto abierto de X. Sin embargo, hay una relación sencilla entre estos conceptos, que estableceremos ahora.

2.30 Teorema *Supongamos $Y \subset X$. Un subconjunto E de Y es abierto relativo en Y si, y solo si $E = Y \cap G$ para algún subconjunto abierto G de X.*

Demostración Supongamos que E es abierto relativo en Y. Para cada $p \in E$ hay un número positivo r_p, tal que las condiciones $d(p,q) < r_p$ y $q \in Y$ implican que $q \in E$. Sea V_p el conjunto de todos los $q \in X$, tales que $d(p,q) < r_p$ y definamos

$$G = \bigcup_{p \in E} V_p .$$

Por los Teoremas 2.19 y 2.24, G será un subconjunto abierto de X.

Como $p \in V_p$ para todo $p \in E$, es claro que $E \subset G \cap Y$.

Por nuestra elección de V_p, tenemos que $V_p \cap Y \subset E$ para todo $p \in E$, de modo que $G \cap Y \subset E$. Así pues, $E = G \cap Y$, con lo que queda demostrada la mitad del teorema.

Inversamente, si G es abierto en X y $E = G \cap Y$, todo $p \in E$ tiene una vecindad $V_p \subset G$. Por tanto, $V_p \cap Y \subset E$, de modo que E es abierto relativo en Y.

CONJUNTOS COMPACTOS

2.31 Definición Llamaremos *cubierta abierta de un conjunto E* en un espacio métrico X a la colección $\{G_\alpha\}$ de subconjuntos abiertos de X, tales que $E \subset \bigcup_\alpha G_\alpha$.

2.32 Definición Se dice que un subconjunto K de un espacio métrico X es *compacto* si toda cubierta abierta de K contiene una subcubierta *finita*.

Más explícitamente, la condición es que si $\{G_\alpha\}$ es una cubierta abierta de K, hay un número finito de índices $\alpha_1, \ldots \alpha_n$, tales que

$$K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n} .$$

La noción de compacticidad es de gran importancia en análisis, especialmente en relación con la continuidad. (Cap. 4).

Se ve claramente que todo conjunto finito es compacto. La existencia de una extensa clase de conjuntos compactos infinitos en R^k, se deduce del Teorema 2.41.

Ya hemos observado (en la Sec. 2.29) que si $E \subset Y \subset X$, E puede ser abierto relativo en Y, sin ser abierto relativo en X. La propiedad de ser abierto depende, pues, del espacio en el que está sumergido E. Igualmente es cierto para la propiedad de ser cerrado.

Sin embargo, es más fácil utilizar la compacticidad, del modo que vamos a ver. Para formular el próximo teorema, diremos provisionalmente, que K es compacto relativo en X si se cumplen las condiciones de la Definición 2.32.

2.33 Teorema *Supongamos $K \subset Y \subset X$. K es compacto relativo en X si, y solo si K es compacto relativo en Y.*

En virtud de este teorema podemos, en muchos casos, considerar conjuntos compactos como espacios métricos en sí mismos, sin prestar atención al espacio que los contiene. En particular, aunque tiene escaso sentido hablar de espacios *abiertos* o *cerrados* (todo espacio métrico X es un subconjunto abierto de sí mismo, así como también un subconjunto cerrado de sí mismo), tendrá sentido hablar de espacios métricos *compactos*.

Demostración supongamos K compacto relativo en X, y sea $\{V_\alpha\}$ una colección de conjuntos, abierta relativa en Y, tal que $K \subset \bigcup_\alpha V_\alpha$. Por el Teorema 2.30, existen conjuntos G_α, abiertos relativos en X, tales que $V_\alpha = Y \cap G_\alpha$, para todo α; y como K es compacto relativo en X, tendremos:

(22) $$K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}$$

para cierta elección de un número finito de índices $\alpha_1, \ldots, \alpha_n$. Como $K \subset Y$ (22) implica que

(23) $$K \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}.$$

Lo que demuestra que K es compacto relativo en Y.

Inversamente, supongamos que K es compacto relativo en Y y sea $\{G_\alpha\}$ una colección de subconjuntos abiertos de X que cubre a K. Hagamos $V_\alpha = Y \cap G_\alpha$: (23) se cumplirá para cierta elección de $\alpha_1, \ldots, \alpha_n$, y como $V_\alpha \subset G_\alpha$ (23) implica (22); lo que completa la demostración.

2.34 Teorema *Los subconjuntos compactos de espacios métricos, son cerrados.*

Demostración Sea K un subconjunto compacto de un espacio métrico X. Demostraremos que el complemento de K es un subconjunto abierto de X.

Supongamos $p \in X$ y $p \notin K$. Si $q \in K$, sean V_q y W_q vecindades de p y q, respectivamente, de radios menores que $\frac{1}{2}d(p,q)$ [ver la Definición 2.18(*a*)]. Como K es compacto, hay un número finito de puntos $q_1, \ldots, q_n$ en K, tales que

$$K \subset W_{q_1} \cup \cdots \cup W_{q_n} = W.$$

Si $V = V_{q_1} \cap \ldots \cap V_{q_n}$, V es una vecindad de p que no interseca a W. Por tanto, $V \subset K^c$, de modo que p es un punto interior de K^c; de donde se deduce el teorema.

2.35 Teorema *Los subconjuntos cerrados de conjuntos compactos, son compactos.*

Demostración Supongamos $F \subset K \subset X$; F cerrado (relativo en X) y K compacto. Sea $\{V_\alpha\}$ una cubierta abierta de F. Si se añade F^c a $\{V_\alpha\}$, se obtiene una cubierta abierta Ω de K. Como K es compacto, hay una subcolección finita Φ de Ω que cubre a K y, por tanto, a F. Si F^c es un elemento de Φ, podemos sacarle de Φ, quedando aun una cubierta abierta de F. Hemos demostrado así que una subcolección finita de $\{V_\alpha\}$ cubre a F.

Corolario *Si F es cerrado y K compacto, $F \cap K$ es compacto*

Demostración Los Teoremas 2.24(*b*) y 2.34, demuestran que $F \cap K$ es cerrado; como $F \cap K \subset K$, el teorema 2.35 demuestra que $F \cap K$ es compacto.

2.36 Teorema Si *$\{K_\alpha\}$ es una colección de subconjuntos compactos de un espacio métrico X, tal que la intersección de toda subcolección finita de $\{K_\alpha\}$ es no vacía, $\bigcap K_\alpha$ es no vacía.*

Demostración Tomemos un elemento K_1 de $\{K_\alpha\}$ y hagamos $G_\alpha = K_\alpha^c$. Admitamos que ningún punto de K_1 pertenece a todos los K_α. En estas condiciones, los conjuntos G_α forman una cubierta abierta de K_1, y como K_1 es compacto, hay un número finito de índices $\alpha_1, \ldots, \alpha_n$, tal que $K_1 \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}$. Pero esto significa que

$$K_1 \cap K_{\alpha_1} \cap \cdots \cap K_{\alpha_n}$$

es vacío, en contra de la hipótesis.

Corolario *Si $\{K_n\}$ es una sucesión de conjuntos compactos no vacíos, tales que $K_n \supset K_{n+1}$ $(n = 1, 2, 3, \ldots)$, $\bigcap_1^\infty K_n$ es no vacío.*

2.37 Teorema *Si E es un subconjunto infinito de un conjunto compacto K, E tiene un punto límite en K.*

Demostración Si ningún punto de K fuera punto límite de E, todo $q \in K$ tendría una vecindad V_q que contendría a lo más un punto de E (esto es, q si $q \in E$). Está claro que ninguna subcolección finita de $\{V_q\}$ puede cubrir a E; y lo mismo sucede con K, luego $E \subset K$, lo que contradice la compacticidad de K.

2.38 Teorema *Si $\{I_n\}$ es una sucesión de intervalos en R^1, tal que $I_n \supset I_{n+1}$ $(n = 1, 2, 3, \ldots)$, $\bigcap_1^\infty I_n$ es no vacía.*

Demostración Si $I_n = [a_n, b_n]$, sea E el conjunto de todos los a_n. E será no vacío y acotado superiormente (por b_1). Sea x el sup de E. Si m y n son enteros positivos:

$$a_n \leq a_{m+n} \leq b_{m+n} \leq b_m,$$

de modo que $x \leq b_m$ para todo m. Como es obvio que $a_m \leq x$, vemos que $x \in I_m$ para $m = 1, 2, 3, \ldots$

2.39 Teorema *Sea k un entero positivo. Si $\{I_n\}$ es una sucesión de celdas-k, tales que $I_n \supset I_{n+1}$ $(n = 1, 2, 3, \ldots)$, $\bigcap_1^\infty I_n$ es no vacío.*

Demostración Supongamos que I_n consta de todos los puntos $\mathbf{x} = (x_1, \ldots, x_k)$, tales que

$$a_{n,j} \leq x_j \leq b_{n,j} \qquad (1 \leq j \leq k; n = 1, 2, 3, \ldots),$$

y hagamos $I_{n,j} = [a_{n,j}, b_{n,j}]$. Para cada j, la sucesión $\{I_{n,j}\}$ satisface la hipótesis del Teorema 2.38. Por tanto, existen números reales x_j^* $(1 \leq j \leq k)$, para los cuales

$$a_{n,j} \leq x_j^* \leq b_{n,j} \qquad (1 \leq j \leq k; n = 1, 2, 3, \ldots).$$

Llamando $\mathbf{x}^* = (x_1^*, \ldots, x_k^*)$, vemos que $\mathbf{x}^* \in I_n$ para $n = 1, 2, 3, \ldots$ de lo que se deduce el teorema.

2.40 Teorema *Toda celda-k es compacta.*

Demostración Sea I una celda-k constituida por todos los puntos $\mathbf{x} = (x_1, \ldots, x_k)$, tales que $a_j \leq x_j \leq b_j$ $(1 \leq j \leq k)$. Hagamos

$$\delta = \left\{\sum_1^k (b_j - a_j)^2\right\}^{1/2}$$

Será $|\mathbf{x} - \mathbf{y}| \leq \delta$, si $\mathbf{x} \in I$, e $\mathbf{y} \in I$.

Supongamos, para llegar a una contradicción, que existe una cubierta abierta $\{G_\alpha\}$ de I que no contiene ninguna subcubierta finita de I. Hagamos $c_j = (a_j + b_j)/2$. Los intervalos $[a_j, c_j]$ y $[c_j, b_j]$ determinarán 2^k celdas-k Q_i, cuya unión es I. Al menos uno de estos conjuntos Q_i, que llamamos I_1, no puede ser cubierto por ninguna subcolección finita de $\{G_\alpha\}$ (de otro modo I podría ser así cubierto). A continuación, dividiremos I_1 y continuaremos el proceso, obteniendo así una sucesión $\{I_n\}$ con las siguientes propiedades

(*a*) $I \supset I_1 \supset I_2 \supset I_3 \supset \cdots$;
(*b*) I_n no es cubierto por ninguna subcolección finita de $\{G_\alpha\}$
(*c*) Si $\mathbf{x} \in I_n$ y $\mathbf{y} \in I_n$, $|\mathbf{x} - \mathbf{y}| \leq 2^{-n}\delta$.

Por (*a*) y el Teorema 2.39, hay un punto $\mathbf{x}^*$ que pertenece a todo I_n. Para algún α, $\mathbf{x}^* \in G_\alpha$. Como G_α es abierto, existe un $r > 0$, tal que $|\mathbf{y}-\mathbf{x}^*| < r$ implica que $\mathbf{y} \in G_\alpha$. Si n es tan grande que $2^{-n}\delta < r$ (tal n existe, porque de otro modo sería $2^n \leq \delta/r$ para todo entero positivo n, lo que es absurdo) porque R es arquimediano (*c*) implica que $I_n \subset G_\alpha$, lo que está en contradicción con (*b*), quedando completa la demostración.

La equivalencia de (*a*) y (*b*) del próximo teorema, se conoce por teorema de Heine-Borel.

2.41 Teorema *Si un conjunto E en R^k tiene una de las tres propiedades siguientes, tiene también las otras dos.*

(*a*) *E es cerrado y acotado*
(*b*) *E es compacto*
(*c*) *Todo subconjunto infinito de E tiene un punto límite en E.*

Demostración Si se cumple (*a*) $E \subset I$ para alguna celda-k I y se deduce (*b*) de los Teoremas 2.40 y 2.35. El Teorema 2.37 demuestra que (*b*) implica (*c*). Queda por demostrar que (*c*) implica (*a*).

Si E no es acotado, contiene puntos $\mathbf{x}_n$ con

$$|\mathbf{x}_n| > n \qquad (n = 1, 2, 3, \ldots).$$

El conjunto S constituido por estos puntos $\mathbf{x}_n$ es infinito y se ve inmediatamente que no tiene ningún punto de límite en R^k y, por tanto, en E, por lo que (*c*) tiene como consecuencia que E es acotado.

Si E no es cerrado, hay un punto $\mathbf{x}_0 \in R^k$ que es punto límite de E, pero no pertenece a E. Para $n = 1, 2, 3, \ldots$ existen puntos $\mathbf{x}_n \in E$, tales que $|\mathbf{x}_n - \mathbf{x}_0| < 1/n$. Sea S el conjunto de estos puntos $\mathbf{x}_n$: S será infinito (de otro modo $|\mathbf{x}_n - \mathbf{x}_0|$ tendría un valor positivo constante para infinitos n) y tiene a $\mathbf{x}_0$ como punto límite y no tiene ningún otro en R^k, porque si $\mathbf{y} \in R^k$ y $\mathbf{y} \neq \mathbf{x}_0$ sería

$$\begin{aligned} |\mathbf{x}_n - \mathbf{y}| &\geq |\mathbf{x}_0 - \mathbf{y}| - |\mathbf{x}_n - \mathbf{x}_0| \\ &\geq |\mathbf{x}_0 - \mathbf{y}| - \frac{1}{n} \geq \frac{1}{2}|\mathbf{x}_0 - \mathbf{y}| \end{aligned}$$

para todos los valores de n, excepto un número finito de ellos, lo que demuestra que $\mathbf{y}$ no es punto límite de S (Teorema 2.20).

Por tanto, S no tiene puntos límite en E; por lo que E debe ser cerrado si se cumple (c).

Debemos hacer notar, en este punto, que (b) y (c) son equivalentes en todo espacio métrico (Ejer. 26), pero que (a), en general, no implica (b) y (c). Se dan ejemplos en el Ejercicio 16 y en el espacio $\mathscr{L}^2$, tratado en el capítulo 11.

2.42 Teorema (Weierstrass) *Todo subconjunto infinito acotado de R^k tiene un punto límite en R^k.*

Demostración Por ser acotado, el conjunto E en cuestión es un subconjunto de una celda-k $I \subset R^k$. Por el Teorema 2.40, I es compacto; y, por tanto, E tiene un punto límite en I, por Teorema 2.37.

CONJUNTOS PERFECTOS

2.43 Teorema *Sea P un conjunto perfecto no vacío en R^k. P es no numerable.*

Demostración Como P tiene puntos límite, debe ser infinito. Supongámosle numerable y representemos sus puntos por $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3, \ldots$ Construiremos una sucesión $\{V_n\}$ de vecindades, como sigue:

Sea V_1 una vecindad de $\mathbf{x}_1$. Si V_1 contiene a todo $\mathbf{y} \in R^k$, tal que $|\mathbf{y} - \mathbf{x}_1| < r$, se define la *cerradura* $\overline{V}_1$ correspondiente, como el conjunto de todos los $\mathbf{y} \in R^k$, tales que $|\mathbf{y} - \mathbf{x}_1| \leq r$.

Supongamos que se ha construido V_n de modo que $V_n \cap P$ es no vacío. Como todo punto de P es punto límite de P, hay una vecindad V_{n+1}, tal que (i) $\overline{V}_{n+1} \subset V_n$; (ii) $\mathbf{x}_n \notin \overline{V}_{n+1}$; (iii) $V_{n+1} \cap P$ es no vacío. Por (iii), V_{n+1} satisface nuestras hipótesis inductivas y puede realizarse la construcción.

Hagamos $K_n = \overline{V}_n \cap P$. Como $\overline{V}_n$ es cerrado y acotado, $\overline{V}_n$ es compacto. Como $\mathbf{x}_n \notin K_{n+1}$ ningún punto de P pertenece a $\bigcap_1^\infty K_n$. Como $K_n \subset P$, esto significa que $\bigcap_1^\infty K_n$ es vacío. Pero todo K_n es no vacío por (iii) y $K_n \supset K_{n+1}$ por (i), lo que contradice al corolario del Teorema 2.36.

Corolario *Todo intervalo $[a,b]$ ($a < b$) es no numerable. En particular, el conjunto de todos los números reales es no numerable.*

2.44 El conjunto de Cantor El conjunto que vamos a construir, ahora demuestra que existen conjuntos perfectos en R^1 que no contienen ningún segmento.

Sea E_0 el intervalo [0,1]. Separemos el segmento $(\frac{1}{3}, \frac{2}{3})$, y sea E_1 la reunión de los intervalos.

$$[0, \tfrac{1}{3}]\ [\tfrac{2}{3}, 1].$$

Separemos los tercios centrales de estos intervalos, y sea E_2 la unión de los intervalos

$$[0, \tfrac{1}{9}], [\tfrac{2}{9}, \tfrac{3}{9}], [\tfrac{6}{9}, \tfrac{7}{9}], [\tfrac{8}{9}, 1].$$

Continuando de este modo; obtenemos una sucesión de conjuntos compactos E_n, tales que

(*a*) $E_1 \supset E_2 \supset E_3 \supset \cdots$;
(*b*) E_n es la unión de 2^n intervalos, cada uno de longitud 3^{-n}.

El conjunto

$$P = \bigcap_{n=1}^{\infty} E_n$$

se llama *conjunto de Cantor*. P es, evidentemente, compacto y el Teorema 2.36 demuestra que es no vacío.

Ningún segmento de la forma

$$\left(\frac{3k+1}{3^m}, \frac{3k+2}{3^m}\right), \tag{24}$$

donde k y m son enteros positivos, tiene un punto común con P. Como todo segmento (α,β) contiene a un segmento de la forma (24), si

$$3^{-m} < \frac{\beta - \alpha}{6},$$

P no contiene a ningún segmento.

Para demostrar que P es perfecto basta hacer ver que no contiene ningún punto aislado. Sea $x \in P$ y S un segmento que contiene a x. Sea I_n el intervalo de E_n que contiene a x. Elijamos n suficientemente grande para que $I_n \subset S$ y sea x_n un extremo de I_n tal que $x_n \neq x$.

De la construcción de P se deduce que $x_n \in P$, por lo que x es un punto límite de P y éste es perfecto.

Una de las propiedades más interesantes del conjunto de Cantor es que nos proporciona un ejemplo de conjunto no numerable de medida cero (el concepto de medida, se estudiará en el Cap. 11).

CONJUNTOS CONEXOS

2.45 Definición Dos subconjuntos A y B de un espacio métrico X se dice que son *separados* si $A \cap \bar{B}$ y $\bar{A} \cap B$ son vacíos, es decir, si ningún punto de A pertenece a la cerradura de B y ningún punto de B pertenece a la cerradura de A.

Un conjunto $E \subset X$ se dice que es *conexo* si E *no* es la unión de dos conjuntos separados que no son vacíos.

2.46 Observación Es obvio que los conjuntos separados son ajenos, pero los conjuntos ajenos no necesariamente son separados. Por ejemplo, el intervalo $[0, 1]$ y el segmento $(1, 2)$ *no* son separados, debido a que 1 es un punto límite de $(1, 2)$. No obstante, los segmentos y $(0, 1)$ son separados.

Los subconjuntos conexos de la recta real tienen una estructura particularmente sencilla:

2.47 Teorema *Un subconjunto E de la recta real R^1 es conexo si, y solamente si tiene la siguiente propiedad: Si $x \in E$, $y \in E$, y $x < z < y$, entonces $z \in E$.*

Demostración Si existe $x \in E$, $y \in E$, y algún $z \in (x,y)$ tal que $z \notin E$, entonces $E = A_z \cup B_z$ donde

$$A_z = E \cap (-\infty, z), \qquad B_z = E \cap (z, \infty).$$

Como $x \in A_z$ y $y \in B_z$, A y B no son vacíos. Ya que $A_z \subset (-\infty, z)$ y $B_z \subset (z, \infty)$, estos son separados. De aquí E no es conexo.

De manera inversa, supóngase que E no es conexo. Entonces hay conjuntos A y B separados que no son vacíos tales que $A \cup B = E$. Si se toman $x \in A$, $y \in B$, y se supone (sin perder generalidad) que $x < y$. Se define ahora

$$z = \sup (A \cap [x, y]).$$

Por el Teorema 2.28, $z \in \bar{A}$; de aquí que $z \notin B$. En particular, $x \leq z < y$.

Si $z \notin A$, se deduce que $x < z < y$ y $z \notin E$.

Si $z \in A$, entonces $z \notin \bar{B}$, en consecuencia existe z_1 tal que $z < z_1 < y$ y $z_1 \notin B$. Por lo tanto, $x < z_1 < y$ y $z_1 \notin E$.

EJERCICIOS

1. Demostrar que el conjunto vacío es un subconjunto de cada conjunto.
2. Se dice que un número complejo z es *algebraico* si hay enteros $a_0, \ldots, a_n$, que no son todos cero, tales que

$$a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n = 0.$$

Demostrar que el conjunto de todos los números algebraicos es numerable. *Sugerencia*: Para cada entero positivo N hay solo un número finito de ecuaciones con

$$n + |a_0| + |a_1| + \cdots + |a_n| = N.$$

3. Demostrar que existen números reales que no son algebraicos.

4. ¿Es numerable el conjunto de todos los números reales irracionales?

5. Construir un conjunto acotado de números reales que tenga exactamente tres puntos límites.

6. Si E' es el conjunto de todos los puntos límite de un conjunto E. Demostrar que E' es cerrado. Demostrar también que E y $\bar{E}$ tienen los mismos puntos límite. (Recordar que $\bar{E} = E \cup E'$.) ¿Siempre tienen E y E' los mismos puntos límite?

7. Sean $A_1, A_2, A_3, \ldots$ subconjuntos de un espacio métrico.
(*a*) Si $B_n = \bigcup_{i=1}^{n} A_i$, demostrar que $\bar{B}_n = \bigcup_{i=1}^{n} \bar{A}_i$, para $n = 1, 2, 3, \ldots$.
(*b*) Si $B = \bigcup_{i=1}^{\infty} A_i$, demostrar que $\bar{B} \supset \bigcup_{i=1}^{\infty} \bar{A}_i$.
Mostrar, con un ejemplo, que esta inclusión puede ser propia.

8. ¿Es cada punto de cada conjunto abierto $E \subset R^2$ un punto límite de E? Responder la misma pregunta para conjuntos cerrados en R^2.

9. Sea E° el conjunto de todos los puntos interiores de un conjunto E. [Véase la Definición 2.18(*e*); a E° se le denomina el *interior* de E.]
(*a*) Demostrar que E° es siempre abierto.
(*b*) Demostrar que E es abierto si y solo si, $E^\circ = E$.
(*c*) Si $G \subset E$ y G es abierto, demostrar que $G \subset E^\circ$.
(*d*) Demostrar que el complemento de E° es la cerradura del complemento de E.
(*e*) ¿Tienen siempre E y $\bar{E}$ los mismos interiores?
(*f*) ¿Tienen siempre E y E° las mismas cerraduras?

10. Sea X un conjunto infinito. Si para $p \in X$ y $q \in X$ se define

$$d(p, q) = \begin{cases} 1 & (\text{si } p \neq q) \\ 0 & (\text{si } p = q). \end{cases}$$

Demostrar que esto es una métrica. ¿Cuáles subconjuntos del espacio métrico resultante son abiertos? ¿Cuáles son cerrados? ¿Cuáles son compactos?

11. Si para $x \in R^1$ y $y \in R^1$ se define

$$d_1(x, y) = (x - y)^2,$$
$$d_2(x, y) = \sqrt{|x - y|},$$
$$d_3(x, y) = |x^2 - y^2|,$$
$$d_4(x, y) = |x - 2y|,$$
$$d_5(x, y) = \frac{|x - y|}{1 + |x - y|}.$$

Determinar cuales de éstas son métricas.

12. Si $K \subset R^1$ consta de 0 y los números $1/n$, para $n = 1, 2, 3, \ldots$. Demostrar que

K es compacto, directamente de la definición (sin usar el teorema de Heine-Borel).

13. Construir un conjunto de números reales compacto en el cual sus puntos límites formen un conjunto numerable.

14. Dar un ejemplo de una cubierta abierta del segmento $(0, 1)$ que no tenga una subcubierta finita.

15. Mostrar que el Teorema 2.36 y su Corolario son falsos (en R^1, por ejemplo) si se reemplaza la palabra "compacto" por "cerrado" o por "acotado".

16. Considerar al conjunto de los números racionales Q, como un espacio métrico con $d(p,q) = |p-q|$. Sea E el conjunto de todos los $p \in Q$ tales que $2 < p^2 < 3$. Mostrar que E es cerrado y acotado en Q, pero que E no es compacto. ¿Es E abierto en Q?

17. Si E es el conjunto de todos los $x \in [0, 1]$ cuya expansión decimal contiene únicamente a los dígitos 4 y 7. ¿Es numerable E? ¿Es denso E en $[0, 1]$? ¿Es E compacto? ¿Es E perfecto?

18. ¿Existe un conjunto perfecto que no es vacío en R^1 que no contiene ningún número racional?

19. (*a*) Si A y B son conjuntos cerrados ajenos en algún espacio métrico X, demostrar que son separados.

(*b*) Demostrar lo mismo para conjuntos ajenos abiertos.

(*c*) Si $p \in X$, $\delta > 0$ son fijos y se define A como el conjunto de todos los $q \in X$ para los cuales $d(p,q) < \delta$ y B se define de manera similar, con $>$ en lugar de $<$. Demostrar que A y B son separados.

(*d*) Demostrar que cada espacio métrico conexo que tiene al menos dos puntos no es numerable. *Sugerencia*: Usar (*c*).

20. ¿Las cerraduras y los interiores de conjuntos conexos son siempre conexos? (Considérense los subconjuntos de R^2.)

21. Sean A y B subconjuntos separados de algún R^k, supóngase además que $\mathbf{a} \in A$, $\mathbf{b} \in B$, y defínase

$$\mathbf{p}(t) = (1-t)\mathbf{a} + t\mathbf{b}$$

para $t \in R^1$. Haciendo $A_0 = \mathbf{p}^{-1}(A)$, $B_0 = \mathbf{p}^{-1}(B)$. [Entonces $t \in A_0$ si y solo si $\mathbf{p}(t) \in A$.]

(*a*) Demostrar que A_0 y B_0 son subconjuntos separados de R^1.

(*b*) Demostrar que existe un $t_0 \in (0, 1)$ tal que $\mathbf{p}(t_0) \notin A \cup B$.

(*c*) Demostrar que cada subconjunto convexo de R^k es conexo.

22. Se dice que un espacio métrico es *separable* si contiene un subconjunto denso numerable. Demostrar que R^k es separable. *Sugerencia*: Considerar el conjunto de los puntos que tienen coordenadas racionales.

23. Se dice que una colección $\{V_\alpha\}$ de subconjuntos abiertos de X es una *base* si se cumplen las condiciones siguientes: Para todo $x \in X$ y cada conjunto abierto $G \subset X$ tal que $x \in G$, tenemos que $x \in V_\alpha \subset G$ para algún α. En otras palabras, todo conjunto abierto en X, es la unión de una subcolección de $\{V_\alpha\}$.

Demostrar que todo espacio métrico separable tiene una base *numerable*.

Sugerencia: Tomar todas las vecindades con radio racional y centro en algún subconjunto denso numerable de X.

24. Sea X un espacio métrico en el cual cada subconjunto infinito tiene un punto límite. Demostrar que X es separable. *Sugerencia*: Fijar $\delta > 0$ y tomar $x_1 \in X$. Elegidos $x_1, \ldots, x_j \in X$, escoger $x_{j+1} \in X$, si es posible, de modo que $d(x_i, x_{j+1}) \geq \delta$ para $i = 1, \ldots, j$. Demostrar que este proceso ha de detenerse después de un número finito de pasos, y que, por tanto, X puede ser cubierto por un número finito de vecindades de radio δ. Tomar $\delta = 1/n (n = 1, 2, 3, \ldots)$ y considerar los centros de las correspondientes vecindades.

25. Demostrar que cada espacio métrico K compacto tiene una base numerable, y que K es por lo tanto separable. *Sugerencia*: Para cada entero positivo n, hay un número finito de vecindades de radio $1/n$ cuya unión cubre a K.

26. Sea X un espacio métrico en el que cada subconjunto infinito tiene un punto límite. Demostrar que X es compacto. *Sugerencia*: Por los ejercicios 23 y 24, X tiene una base numerable. Se deduce que toda cubierta abierta de X tiene una subcubierta *numerable* $\{G_n\}$, $n = 1, 2, 3, \ldots$ Si ninguna subcolección finita de $\{G_n\}$ cubre a X, el complemento F_n de $G_1 \cup \ldots \cup G_n$ es no vacío para todo n, pero $\bigcap F_n$ es vacío. Si E es un conjunto que contiene un punto de cada F_n, considerar un punto límite de E y llegar a una contradicción.

27. Se dice que un punto p en un espacio métrico X es un *punto* de *condensación* de un conjunto $F \subset X$ si cada vecindad de p contiene un número de puntos de E que no es numerable.

    Suponer que $E \subset R^k$ es no numerable, y sea P el conjunto de todos los puntos de condensación de E. Demostrar que P es perfecto y que a lo más un número de puntos numerable de E no está en P. En otras palabras, mostrar que $P^c \cap E$ es a lo más numerable. *Sugerencia:* Sea $\{V_n\}$ una base numerable de R^k, y W la unión de los V_n para los cuales $E \cap V_n$ es a lo más numerable, entonces muéstrese que $P = W^c$.

28. Demostrar que todo conjunto cerrado en un espacio métrico separable, es la unión de un conjunto (es posible que vacío) perfecto y un conjunto que es a lo sumo numerable. (*Corolario:* Todo conjunto cerrado numerable en R^k posee puntos aislados.) *Sugerencia:* Véase el Ejer. 27.

29. Demostrar que todo conjunto abierto en R^1 es la unión de una colección a lo sumo numerable de segmentos ajenos. *Sugerencia:* Utilizar el Ejer. 22.

30. Imitar la demostración del teorema 2.43 para llegar al resultado siguiente:

    Si $R^k = \bigcup_1^\infty F_n$ donde cada F_n es un subconjunto cerrado de R^k, al menos uno de los F_n tiene un interior no vacío.

    Propiedad equivalente: Si G_n es un subconjunto abierto dentro de R^k, para $n = 1, 2, 3, \ldots, \bigcap_1^\infty G_n$ es no vacío (de hecho, es denso en R^k).

    (Es un caso particular del teorema de Baire; ver el Ejer. 22, Cap. 3, para el caso general.)

3

SUCESIONES NUMÉRICAS Y SERIES

Como indica el título, este capítulo tratará principalmente sobre sucesiones y series de números complejos. Sin embargo, los hechos fundamentales sobre convergencia se explican con la misma facilidad para casos más generales. Las tres primeras secciones tratarán, en consecuencia, de sucesiones en espacios euclideanos, o incluso en espacios métricos.

SUCESIONES CONVERGENTES

3.1 Definición Se dice que una sucesión $\{p_n\}$ en un espacio métrico X, *converge* si hay un punto $p \in X$ con las siguientes propiedades: para cada $\varepsilon > 0$, existe un número entero N tal que $n \geq N$ implica que $d(p_n, p) < \varepsilon$. (d representa la distancia en X.)

En este caso, decimos también que $\{p_n\}$ converge hacia p, o que p es el límite de $\{p_n\}$ [ver el Teorema 3.2(b)] y escribimos $p_n \to p$, o

$$\lim_{n \to \infty} p_n = p.$$

Si $\{p_n\}$ no converge, se dice que *diverge*.

Puede ser conveniente hacer resaltar que nuestra definición de «sucesión convergente» depende no solamente de $\{p_n\}$, sino también de X; por ejemplo, la sucesión $\{1/n\}$ converge en R^1 (hacia 0), pero no lo hace en el

conjunto de los números reales positivos [con $d(x,y) = |x - y|$]. En casos de posible ambigüedad, debemos ser más precisos y especificar «convergente en X», mejor que solamente «convergente».

Recordemos que el conjunto de todos los puntos p_n ($n - 1, 2, 3, \ldots$) es el rango de $\{p_n\}$. El rango de una sucesión puede ser un conjunto finito, o puede ser infinito. Se dice que la sucesión $\{p_n\}$ es *acotada* si lo es su rango.

Como ejemplo, consideremos las siguientes sucesiones de números complejos (esto es, $X = R^2$).

(*a*) Si $s_n = 1/n$, es $\lim\limits_{n\to\infty} s_n = 0$; el rango es infinito y la sucesión es acotada.
(*b*) Si $s_n = n^2$, la sucesión $\{s_n\}$ no es acotada, es divergente y tiene un rango infinito.
(*c*) Si $s_n = 1 + [(-1)^n/n]$, la sucesión $\{s_n\}$ converge hacia 1, es acotada y tiene rango infinito.
(*d*) Si $s_n = i^n$, la sucesión $\{s_n\}$ es divergente, es acotada, y tiene un rango finito.
(*e*) Si $s_n = 1$ ($n = 1, 2, 3\ldots$), $\{s_n\}$ converge hacia 1, es acotado y tiene un rango finito.

Resumiremos a continuación algunas propiedades importantes de las sucesiones convergentes en espacios métricos.

3.2 Teorema *Sea $\{p_n\}$ una sucesión en un espacio métrico X.*

(*a*) *$\{p_n\}$ converge hacia $p \in X$ si, y solo si toda vecindad de p contiene todos los términos de $\{p_n\}$, salvo un número finito de ellos.*
(*b*) *Si $p \in X$; $p' \in X$ y $\{p_n\}$ convergen hacia p y hacia p' entonces $p' = p$.*
(*c*) *Si $\{p_n\}$ converge, es acotada.*
(*d*) *Si $E \subset X$ y p es un punto de límite de E, existe una sucesión $\{p_n\}$ en E para la cual $p = \lim\limits_{n\to\infty} p_n$.*

Demostración (*a*) Supongamos $p_n \to p$ y sea V una vecindad de p. Para algún $\varepsilon > 0$, las condiciones $d(p,q) < \varepsilon$ y $q \in X$ implican que $q \in V$. En correspondencia con este ε, existe un N tal que $n \geq N$ implica que $d(p_n,p) < \varepsilon$. Por tanto, $n \geq N$ implica que $p_n \in V$.

Inversamente, supongamos que toda vecindad de p contiene a todos los p_n, salvo un número finito. Tomemos $\varepsilon > 0$, y sea V el conjunto de todos los $q \in X$ tales que $d(p,q) < \varepsilon$. Por hipótesis, existe un N (correspondiente a este V) tal que $p_n \in V$ si $n \geq N$. Así pues, $d(p_n,p) < \varepsilon$ si $n \geq N$ y, por tanto $p_n \to p$.

(*b*) Sea un $\varepsilon > 0$, dado. Existen dos enteros N, N' tales que

$$n \geq N \quad \text{implica} \quad d(p_n, p) < \frac{\varepsilon}{2},$$

$$n \geq N' \quad \text{implica} \quad d(p_n, p') < \frac{\varepsilon}{2}.$$

De aquí, si $n \geq$ máx (N,N'), tenemos

$$d(p, p') \leq d(p, p_n) + d(p_n, p') < \varepsilon.$$

Como ε era arbitrario, se deduce que $d(p,q') = 0$.

(*c*) Supongamos que $p_n \to p$. Existe un entero N tal que $n > N$ implica que $d(p_n,p) < 1$. Hagamos

$$r = \text{máx}\,\{1, d(p_1, p), \ldots, d(p_N, p)\}.$$

Entonces, $d(p_n,p) \leq r$ para $n = 1, 2, 3, \ldots$

(*d*) Para todo entero positivo n, hay un punto $p_n \in E$, tal que $d(p_n,p) < 1/n$. Dado $\varepsilon > 0$, elijamos N de modo que $N\varepsilon > 1$. Si $n > N$, se deduce que $d(p_n,p) < \varepsilon$. Por tanto, $p_n \to p$, lo que completa la demostración.

Podemos estudiar la relación entre la convergencia y las operaciones algebraicas, para las sucesiones en R^k. Consideremos primeramente sucesiones de números complejos.

3.3 Teorema *Supongamos que* $\{s_n\}$, $\{t_n\}$ *son sucesiones complejas y* $\lim_{n\to\infty} s_n = s$, $\lim_{n\to\infty} t_n = t$. *Se verificará*

(*a*) $\lim_{n\to\infty} (s_n + t_n) = s + t$;

(*b*) $\lim_{n\to\infty} cs_n = cs$, $\lim_{n\to\infty} (c + s_n) = c + s$, *para cualquier número* c;

(*c*) $\lim_{n\to\infty} s_n t_n = st$;

(*d*) $\lim_{n\to\infty} \frac{1}{s_n} = \frac{1}{s}$, *siempre que* $s_n \neq 0$ $(n = 1, 2, 3, \ldots)$, y $s \neq 0$.

Demostración

(*a*) Dado $\varepsilon > 0$, existen dos enteros N_1, N_2 tales que

$$n \geq N_1 \quad \text{implica} \quad |s_n - s| < \frac{\varepsilon}{2},$$

$$n \geq N_2 \quad \text{implica} \quad |t_n - t| < \frac{\varepsilon}{2}.$$

Si $N = \text{máx}\,(N_1, N_2)$, $n \geq N$ implica

$$|(s_n + t_n) - (s + t)| \leq |s_n - s| + |t_n - t| < \varepsilon.$$

Lo que demuestra (a). La demostración de (b) es elemental.

(c) Utilizaremos la identidad

$$s_n t_n - st = (s_n - s)(t_n - t) + s(t_n - t) + t(s_n - s). \tag{1}$$

Dado $\varepsilon > 0$, existen dos enteros N_1, N_2 tales que

$$n \geq N_1 \quad \text{implica} \quad |s_n - s| < \sqrt{\varepsilon},$$
$$n \geq N_2 \quad \text{implica} \quad |t_n - t| < \sqrt{\varepsilon}.$$

Si tomamos $N = \text{máx}\,(N_1, N_2)$, $n \geq N$ implica

$$|(s_n - s)(t_n - t)| < \varepsilon,$$

de modo que

$$\lim_{n \to \infty} (s_n - s)(t_n - t) = 0.$$

Apliquemos (a) y (b) a (1). Se deduce que

$$\lim_{n \to \infty} (s_n t_n - st) = 0.$$

(d) Eligiendo m de modo que $|s_n - s| < \frac{1}{2}|s|$, si $n \geq m$, vemos que

$$|s_n| > \tfrac{1}{2}|s| \qquad (n \geq m).$$

Dado $\varepsilon > 0$, existe un entero $N > m$ tal que $n \geq N$ implica que

$$|s_n - s| < \tfrac{1}{2}|s|^2 \varepsilon.$$

Por tanto, para $n \geq N$,

$$\left|\frac{1}{s_n} - \frac{1}{s}\right| = \left|\frac{s_n - s}{s_n s}\right| < \frac{2}{|s|^2}|s_n - s| < \varepsilon.$$

3.4 Teorema

(a) *Supongamos* $\mathbf{x}_n \in R^k$ ($n = 1, 2, 3, \ldots$) *y*

$$\mathbf{x}_n = (\alpha_{1,n}, \ldots, \alpha_{k,n}).$$

$\{\mathbf{x}_n\}$ *converge hacia* $\mathbf{x} = (\alpha_1, \ldots, \alpha_k)$ *si, y solo si*

(2) $$\lim_{n\to\infty} \alpha_{j,n} = \alpha_j \qquad (1 \le j \le k).$$

(*b*) *Supongamos que* $\{\mathbf{x}_n\}$ *e* $\{\mathbf{y}_n\}$ *son sucesiones en* R^k, $\{\beta_n\}$ *es una sucesión de números reales, y* $\mathbf{x}_n \to \mathbf{x}$; $\mathbf{y}_n \to \mathbf{y}$; $\beta_n \to \beta$. *Tendremos que*

$$\lim_{n\to\infty} (\mathbf{x}_n + \mathbf{y}_n) = \mathbf{x} + \mathbf{y}, \qquad \lim_{n\to\infty} \mathbf{x}_n \cdot \mathbf{y}_n = \mathbf{x} \cdot \mathbf{y}, \qquad \lim_{n\to\infty} \beta_n \mathbf{x}_n = \beta \mathbf{x}.$$

Demostración

(*a*) Si $\mathbf{x}_n \to \mathbf{x}$, las desigualdades

$$|\alpha_{j,n} - \alpha_j| \le |\mathbf{x}_n - \mathbf{x}|,$$

que se deducen inmediatamente de la definición de la norma de R^k demuestran que se cumple (2).

Inversamente, si se cumple (2), a cada $\varepsilon > 0$ corresponde un entero N tal que $n \ge N$ implica que

$$|\alpha_{j,n} - \alpha_j| < \frac{\varepsilon}{\sqrt{k}} \qquad (1 \le j \le k).$$

y por tanto, $n \ge N$ implica que

$$|\mathbf{x}_n - \mathbf{x}| = \left\{\sum_{j=1}^{k} |\alpha_{j,n} - \alpha_j|^2\right\}^{1/2} < \varepsilon,$$

por lo que, $\mathbf{x}_n \to \mathbf{x}$, lo que demuestra (*a*).

El apartado (*b*) se deduce de (*a*) y del Teorema 3.3.

SUBSUCESIONES

3.5 Definición Dada una sucesión $\{p_n\}$, consideremos otra $\{n_k\}$ constituida por enteros positivos, de modo que $n_1 < n_2 < n_3 < \cdots$. La sucesión $\{p_{n_i}\}$ se llama *subsucesión* de $\{p_n\}$. Si $\{p_{n_i}\}$ converge, su límite se llama *límite subsecuencial* de $\{p_n\}$.

Es claro que $\{p_n\}$ converge hacia p si, y solo si toda subsucesión de $\{p_n\}$ converge hacia p. Dejamos los detalles de la demostración al lector.

3.6 Teorema

(*a*) *Si* $\{p_n\}$ *es una sucesión en un espacio métrico compacto* X, *entonces alguna subsucesión de* $\{p_n\}$ *converge hacia un punto de* X.

(*b*) *Toda sucesión acotada en R^k contiene una subsucesión convergente.*

Demostración

(*a*) Sea E el rango de $\{p_n\}$. Si E es finito entonces hay un $p \in E$ y una sucesión $\{n_i\}$ con $n_1 < n_2 < n_3 < \cdots$, tales que

$$p_{n_1} = p_{n_2} = \cdots = p.$$

La subsucesión $\{p_{n_i}\}$ así obtenida converge evidentemente hacia p.

Si E es infinito, el Teorema 2.37 muestra que E tiene un punto límite $p \in X$. Si se escoge n_1 de tal forma que $d(p,p_{n1}) < 1$. Después de escoger $n_1,\ldots, n_{i-1}$, el Teorema 2.20 indica que hay un entero $n_i > n_{i-1}$ tal que $d(p,p_{ni}) < 1/i$. Por lo tanto $\{p_{ni}\}$ converge hacia p.

(*b*) Esto se concluye de (*a*), porque el Teorema 2.41 implica que cada subconjunto acotado de R^k está en un subconjunto compacto de R^k.

3.7 Teorema *Los límites subsecuenciales de una sucesión $\{p_n\}$ en un espacio métrico X forman un subconjunto cerrado de X.*

Demostración Sea E^* el conjunto de los límites subsecuenciales de $\{p_n\}$ y que q un punto límite de E^*. Se tiene que mostrar que $q \in E^*$.

Si se escoge n_1 de manera que $p_{n1} \neq q$. (Si tal n_1 no existe, entonces E^* tiene solo un punto, y no hay nada que demostrar.) Hágase $\delta = d(q,p_{n1})$ y supóngase que se han escogido $n_1,\ldots, n_{i-1}$. Como q es un punto límite de E^*, entonces hay un $x \in E^*$ con $d(x,q) < 2^{-i}\delta$. Ya que $x \in E^*$, entonces hay un $n_i > n_{i-1}$ tal que $d(x,p_{ni}) < 2^{-i}\delta$. Por lo tanto

$$d(q, p_{n_i}) \leq 2^{1-i}\delta$$

para $i = 1, 2, 3,\ldots$. Esto quiere decir que $\{p_{ni}\}$ converge hacia q. En consecuencia $q \in E^*$.

SUCESIONES DE CAUCHY

3.8 Definición Se dice que una sucesión $\{p_n\}$ en un espacio métrico X es una *sucesión de Cauchy* si para todo $\varepsilon > 0$, hay un entero N tal que $d(p_n,p_m) < \varepsilon$ si $n \geq N$ y $m \geq N$.

En el estudio de las sucesiones de Cauchy, del mismo modo que en otros casos que encontraremos más adelante, será útil el siguiente concepto geométrico.

3.9 Definición Sea E un subconjunto de un espacio métrico X, y S el conjunto de todos los números reales de la forma $d(p,q)$ con $p \in E$ y $q \in E$. El sup de S se llama *díametro* de E.

Si $\{p_n\}$ es una sucesión en X y E_N está constituido por los puntos p_N, $p_{N+1}, p_{N+2}, \ldots$, se ve fácilmente, como consecuencia de las dos definiciones anteriores, que $\{p_n\}$ *es una sucesión de Cauchy si, y solo si*

$$\lim_{N\to\infty} \text{diám } E_N = 0.$$

3.10 Teorema

(*a*) *Si $\bar{E}$ es la cerradura de un conjunto E en un espacio métrico X, será*:

$$\text{diám } \bar{E} = \text{diám } E.$$

(*b*) *Si K_n es una sucesión de conjuntos compactos en X, tales que $K_n \supset K_{n+1} (n = 1, 2, 3, \ldots)$ y si*

$$\lim_{n\to\infty} \text{diám } K_n = 0,$$

$\bigcap_1^\infty K_n$ *está constituido por un punto*

Demostración

(*a*) Como $E \subset \bar{E}$, es claro que

$$\text{diám } E \leq \text{diám } \bar{E}.$$

Fijemos $\varepsilon > 0$, y elijamos $p \in \bar{E}$ y $q \in \bar{E}$. Por la definición de $\bar{E}$, existen puntos p', q', en E tales que $d(p,p') < \varepsilon$, y $d(q,q') < \varepsilon$. De aquí,

$$\begin{aligned} d(p;q) &\leq d(p, p') + d(p' q') + d(q', q) \\ &< 2\varepsilon + d(p', q') \leq 2\varepsilon + \text{diám } E. \end{aligned}$$

Se deduce que

$$\text{diám } \bar{E} \leq 2\varepsilon + \text{diám } E,$$

y como ε era arbitrario, queda demostrado (*a*).

(*b*) Hagamos $K = \bigcap_1^\infty K_n$. Por el Teorema 2.36, K es no vacío. Si contiene más de un punto, diám $K > 0$. Pero para todo n, $K_n \supset K$, de modo que diám $K_n \geq$ diám K, lo que contradice la hipótesis de ser diám $K_n \to 0$.

3.11 Teorema

(*a*) *En cualquier espacio métrico X, toda sucesión convergente es una sucesión de Cauchy.*
(*b*) *Si X es un espacio métrico compacto y* $\{p_n\}$ *es una sucesión de Cauchy en X, entonces* $\{p_N\}$ *converge hacia algún punto de X.*
(*c*) *En* R^k *toda sucesión de Cauchy converge.*

Nota: La diferencia entre la definición de convergencia y la de sucesión de Cauchy, es que el límite está incluido explícitamente en la primera y no en la segunda. Así el Teorema 3.11(*b*) permitirá averiguar si una sucesión dada converge o no, sin conocer el límite a que pueda tender.

El hecho (contenido en el Teorema 3.11) de converger una sucesión en R^k si, y solo si es de Cauchy, se llama comúnmente *criterio de convergencia de Cauchy.*

Demostración

(*a*) Si $p_n \to p$ y $\varepsilon > 0$, hay un entero N tal que $d(p,q_n) < \varepsilon$ para todo $n \geq N$. En consecuencia

$$d(p_n, p_m) \leq d(p_n, p) + d(p, p_m) < 2\varepsilon$$

si $n \geq N$ y $m \geq N$. Por lo tanto $\{p_n\}$ es una sucesión de Cauchy.

(*b*) Sea $\{p_n\}$ una sucesión de Cauchy en el conjunto compacto X. Para $N = 1, 2, 3, \ldots$, sea E_N el conjunto que consta de p_N, p_{N+1}, $p_{N+2}, \ldots$. Entonces

(3) $$\lim_{N\to\infty} \text{diám } \bar{E}_N = 0,$$

debido a la Definición 3.9 y al Teorema 3.10(*a*). Cada X es compacto porque es un subconjunto cerrado de un espacio compacto $\bar{E}_N$ (Teorema 2.35). También $E_N \supset E_{N+1}$, así que $\bar{E}_N \supset \bar{E}_{N+1}$.

El Teorema 3.10(*b*) indica que hay un $p \in X$ único, que está en cada $\bar{E}_N$.

Dado $\varepsilon > 0$. Por (3) hay un entero N_0 tal que $\bar{E}_N < \varepsilon$ si $N \geq N_0$. Como $p \in \bar{E}_N$, se deduce que $d(p,q) < \varepsilon$ para cada $q \in \bar{E}_N$, de aquí que para cada $q \in E_N$. En otras palabras $d(p,p_n) < \varepsilon$ si $n \geq N_0$. Esto expresa exactamente que $p_n \to p$.

(*c*) Sea x_n una sucesión de Cauchy en R^k. Defínase E_N como en (*b*) con $\mathbf{x}_i$ en lugar de p_i. Para algún N, diám $E_N < 1$. El rango de $\{\mathbf{x}_n\}$ es

la unión de E_N y el conjunto finito $\{\mathbf{x}_1, \ldots, \mathbf{x}_{N-1}\}$. Por esto $\{\mathbf{x}_n\}$ es acotado. Finalmente de (*b*) se deduce (*c*), debido a que todo subconjunto acotado de R^k tiene cerradura compacta en R^k (Teorema 2.41).

3.12 Definición Si toda sucesión de Cauchy converge en un espacio métrico, se dice que este es *completo.*

Entonces el Teorema 3.11 expresa que *todos los espacios métricos compactos y los espacios euclidianos son completos.* El Teorema 3.11 implica también que *todo subconjunto cerrado E de un espacio métrico completo X es completo.* (Toda sucesión de Cauchy en E es una sucesión de Cauchy en X, por esto converge hacia algún $p \in X$, y ciertamente $p \in E$ porque E es cerrado.) Un ejemplo de un espacio métrico que no es completo es el espacio de todos los números racionales, $d(x,y) = |x - y|$.

El Teorema 3.2(*c*) y el ejemplo (*d*) de la Definición 3.1, demuestran que las sucesiones convergentes son acotadas, pero que las sucesiones acotadas, en R^k no son necesariamente convergentes. Sin embargo, hay un caso importante en el que convergencia *es* equivalente a acotación; esto sucede con las sucesiones monótonas en R^1.

3.13 Definición Se dice que una sucesión $\{s_n\}$ de números reales es:

(*a*) *monótona creciente* si $s_n \le s_{n+1} (n = 1, 2, 3, \ldots)$;
(*b*) *monótona decreciente* si $s_n \ge s_{n+1} (n = 1, 2, 3, \ldots)$.

La clase de las sucesiones monótonas está constituida por las sucesiones crecientes y las decrecientes.

3.14 Teorema *Supongamos que* $\{s_n\}$ *es monótona.* $\{s_n\}$ *converge si, y solo si es acotada.*

Demostración Supongamos que $s_n \le s_{n+1}$ (la demostración es análoga en el otro caso). Sea E el rango de $\{s_n\}$. Si $\{s_n\}$ es acotada, sea s el sup de E. Será

$$s_n \le s \qquad (n = 1, 2, 3, \ldots).$$

Para todo $\varepsilon > 0$, existe un entero N, tal que

$$s - \varepsilon < s_N \le s,$$

porque de otro modo $s - \varepsilon$ sería una cota superior de E. Como $\{s_n\}$ es creciente, $n \ge N$ implica que

$$s - \varepsilon < s_n \le s,$$

lo que demuestra que $\{s_n\}$ converge (hacia s).

La inversa se deduce del Teorema 3.2(c).

LÍMITES SUPERIOR E INFERIOR

3.15 Definición Sea $\{s_n\}$ una sucesión de números reales con las siguientes propiedades: para cada número real M hay un entero N, tal que $n \geq N$ implica que $s_n > M$. Escribimos

$$s_n \to +\infty.$$

Del mismo modo, si para todo número real M existe un entero N tal que $n \geq N$ implica que $s_n \leq M$, escribimos

$$s_n \to -\infty.$$

Se habrá observado que hemos utilizado el símbolo $\to$ (introducido en la Definición 3.1) pra ciertos tipos de sucesiones divergentes, igual que para las convergentes, pero sin que se hayan cambiado las definiciones de convergencia y de límite, dadas en la Definición 3.1.

3.16 Definición Sea $\{s_n\}$ una sucesión de números reales, y E el conjunto de los números x (en el sisten a extendido de números reales) tales que $s_{n_k} \to x$ para alguna subsucesión $\{s_{n_k}$ Este conjunto E contiene todos los límites subsecuenciales definidos en la Definición 3.5 más, posiblemente, los números $+\infty$ y $-\infty$.

Recordemos, ahora, las Definiciones 1.8 y 1.23 y hagamos

$$s^* = \sup E,$$
$$s_* = \inf E.$$

Los números s^* y s_* se llaman *límites superior e inferior* de $\{s_n\}$; utilizaremos la notación

$$\lim_{n\to\infty} \sup s_n = s^*, \qquad \lim_{n\to\infty} \inf s_n = s_*.$$

3.17 Teorema *Sea $\{s_n\}$ una sucesión de números reales, y consideremos que E y s^* tienen el mismo significado que en la Definición 3.16. En estas condiciones, s^* tiene las dos propiedades siguientes:*

(a) $s^* \in E$.
(b) *Si $x > s^*$, existe un entero N tal que $n \geq N$ implica que $s_n < x$.*

Además, s^* es el único número que posee las propiedades (a) y (b).

Para s_* se obtiene, evidentemente, un resultado análogo.

Demostración

(*a*) Si $s^* = +\infty$, E no está acotado superiormente, por lo que no lo está $\{s_n\}$ y existe una subsucesión $\{s_{n_k}\}$ tal que $s_{n_k} \to +\infty$.

Si s^* es real, E está acotado superiormente y existe al menos un límite subsecuencial, de modo que (*a*) se deduce de los Teoremas 3.7 y 2.28.

Si $s^* = -\infty$, E contiene solo un elemento, es decir $-\infty$, y no existe límite subsecuencial. Por tanto, para todo número real M, $s_n > M$ para, a lo sumo, un número finito de valores de n, de modo que $s_n \to -\infty$.

Queda así demostrado (*a*) para todos los casos posibles.

(*b*) Para demostrar (*b*), supongamos que existe un número $x > s^*$ tal que $s_n \geq x$ para infinitos valores de n. En este caso, existe un número $y \in E$ tal que $y \geq x \geq s^*$, en contradicción con la definición de s^*.

Por tanto, s^* satisface a (*a*) y (*b*).

Para demostrar la unicidad, supongamos que existen dos números p y q que satisfacen (*a*) y (*b*), y sea $p < q$. Elijamos x de modo que $p < x < q$. Como p satisface (*b*), tenemos que $s_n < x$ para $n \geq N$. Pero en este caso, q no puede satisfacer a (*a*).

3.18 Ejemplos

(*a*) Sea $\{s_n\}$ una sucesión que contiene a todos los números racionales. Todo número real es un límite subsecuencial y

$$\limsup_{n\to\infty} s_n = +\infty, \qquad \liminf_{n\to\infty} s_n = -\infty.$$

(*b*) Sea $s_n = (-1)^n[1 + (1/n)]$. Será

$$\limsup_{n\to\infty} s_n = 1, \qquad \liminf_{n\to\infty} s_n = -1.$$

(*c*) Para una sucesión $\{s_n\}$ de valores reales, $\lim_{n\to\infty} s_n = s$, si y solo si

$$\limsup_{n\to\infty} s_n = \liminf_{n\to\infty} s_n = s.$$

Terminaremos esta sección con un teorema de gran utilidad, cuya demostración es elemental.

3.19 Teorema *Si $s_n \leq t_n$ para $n \geq N$, con N, prefijado, se verifica que*

$$\lim_{n\to\infty} \inf s_n \le \lim_{n\to\infty} \inf t_n,$$

$$\lim_{n\to\infty} \sup s_n \le \lim_{n\to\infty} \sup t_n.$$

ALGUNAS SUCESIONES ESPECIALES

Calcularemos, ahora, los límites de algunas sucesiones que se presentan frecuentemente. Las demostraciones están basadas todas ellas en la siguiente observación: si $0 \le x_n \le$ para $n \ge N$, siendo N un número dado, y si $s_n \to 0$, se verifica que $x_n \to 0$.

3.20 Teorema

(*a*) *Si* $p > 0$, $\lim_{n\to\infty} \frac{1}{n^p} = 0$.

(*b*) *Si* $p > 0$, $\lim_{n\to\infty} \sqrt[n]{p} = 1$.

(*c*) $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

(*d*) *Si* $p > 0$ *y* α *es real,* $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$.

(*e*) *Si* $|x| < 1$, $\lim_{n\to\infty} x^n = 0$.

Demostración

(*a*) Tomemos $n > (1/\varepsilon)^{1/p}$. (Nótese que aquí se ha usado la propiedad arquimidiana del sistema de los números reales.)
(*b*) Si $p > 1$, hagamos $x_n = \sqrt[n]{p} - 1$. Será $x_n > 0$ y por la fórmula del binomio

$$1 + nx_n \le (1 + x_n)^n = p,$$

de modo que

$$0 < x_n \le \frac{p-1}{n}.$$

Por tanto, $x_n \to 0$. Si $p = 1$, (*b*) es evidente, y si $0 < p < 1$, se obtiene el resultado tomando los recíprocos.
(*c*) Hagamos $x_n = \sqrt[n]{n} - 1$. Será x_n 0, y por el teorema del binomio

$$n = (1 + x_n)^n \ge \frac{n(n-1)}{2} x_n^2.$$

Por tanto

$$0 \leq x_n \leq \sqrt{\frac{2}{n-1}} \qquad (n \geq 2).$$

(*d*) Sea k un entero tal que $k > \alpha$; $k > 0$. Para $n > 2k$.

$$(1+p)^n > \binom{n}{k} p^k = \frac{n(n-1)\cdots(n-k+1)}{k!} p^k > \frac{n^k p^k}{2^k k!}.$$

y de aquí

$$0 < \frac{n^\alpha}{(1+p)^n} < \frac{2^k k!}{p^k} n^{\alpha-k} \qquad (n > 2k).$$

Como $\alpha - k < 0$, $n^{\alpha-k} \to 0$ por (*a*).
(*e*) Hagamos $\alpha = 0$ en (*d*).

SERIES

En lo que resta de este capítulo, todas las sucesiones y series que consideremos tendrán valores complejos, excepto cuando se indique lo contrario explícitamente. En el Ejercicio 15 se hace referencia a la extensión de alguno de los teoremas que siguen, a series con términos en R^k.

3.21 Definición Dada una sucesión $\{a_n\}$ usaremos la notación

$$\sum_{n=p}^{q} a_n \qquad (p \leq q)$$

para expresar la suma $a_p + a_{p+1} + \cdots + a_q$. A $\{a_n\}$ le asociamos una sucesión $\{s_n\}$, donde

$$s_n = \sum_{k=1}^{n} a_k.$$

También utilizaremos para $\{s_n\}$ la expresión simbólica

$$a_1 + a_2 + a_3 + \cdots$$

o, más brevemente

(4) $$\sum_{n=1}^{\infty} a_n.$$

Al símbolo (4) le llamaremos *serie infinita* o solamente *serie*. A los números s_n se le llama *sumas parciales* de la serie. Si $\{s_n\}$ converge hacia s, diremos que la serie *converge* y escribiremos

$$\sum_{n=1}^{\infty} a_n = s.$$

Al número s se le llama suma de la serie; pero debe entenderse claramente que *s es el límite de una sucesión de sumas* y que no se obtiene simplemente por adición.

Si $\{s_n\}$ diverge, se dice que la serie diverge.

A veces, por conveniencia de notación, consideraremos series en la forma

$$\sum_{n=0}^{\infty} a_n . \tag{5}$$

y frecuentemente, cuando no haya peligro de ambigüedad, o si la distinción es de escasa importancia, escribiremos simplemente Σa_n en lugar de (4) o (5).

Es evidente que todo teorema sobre sucesiones, se puede enunciar refiriéndose a las series (haciendo $a_1 = s_1$ y $a_n = s_n - s_{n-1}$ para $n > 1$) y viceversa, pero es, sin embargo, útil considerar ambos conceptos.

El criterio de Cauchy (Teorema 3.11) puede ser enunciado de nuevo en la forma siguiente:

3.22 Teorema *Σa_n converge si, y solo si para cada $\varepsilon > 0$ existe un entero N, tal que*

$$\left|\sum_{k=n}^{m} a_k\right| \leq \varepsilon \tag{6}$$

si $m \geq n \geq N$.

En particular, tomando $m = n$, (6) se convierte en

$$|a_n| \leq \varepsilon \qquad (n \geq N).$$

En otras palabras:

3.23 Teorema Si *Σa_n converge,* $\lim\limits_{n\to\infty} a_n = 0$.

La condición $a_n \to 0$, no es, sin embargo, suficiente para poder asegurar la convergencia de Σa_n. Por ejemplo, la serie

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverge. Para demostrarlo, considérese el Teorema 3.28.

El Teorema 3.14 sobre sucesiones monótonas, tiene también una equivalencia inmediata para las series.

3.24 Teorema *Una serie de términos no negativos converge*[1] *si, y solo si sus sumas parciales forman una sucesión acotada.*

Veamos ahora otro método distinto de estudiar la convergencia, el llamado «criterio de comparación».

3.25 Teorema

(*a*) *Si* $|a_n| \leq c_n$ para $n \geq N_0$, *donde* N_0 *es un entero dado, y* Σc_n *converge, también converge* Σa_n.
(*b*) *Si* $a_n \geq d_n \geq 0$ *para* $n \geq N_0$ *y* Σd_n *diverge, también diverge* Σa_n.

Nótese que (*b*) se aplica solamente a series de términos a_n no negativos.

Demostración Dado $\varepsilon > 0$, existe $N \geq N_0$, tal que $m \geq n \geq N$ implica que

$$\sum_{k=n}^{m} c_k \leq \varepsilon,$$

por el criterio de Cauchy. De aquí, que

$$\left|\sum_{k=n}^{m} a_k\right| \leq \sum_{k=n}^{m} |a_k| \leq \sum_{k=n}^{m} c_k \leq \varepsilon,$$

y se deduce (*a*).

Además (*b*) se deduce de (*a*), porque si Σa_n converge, igual debe suceder con Σd_n [obsérvese que (*b*) se deduce también del Teorema 3.24].

El criterio de comparación es muy útil; para usarlo convenientemente, debemos familiarizarnos con cierto número de series de términos no negativos, cuya convergencia o divergencia es conocida.

SERIES DE TÉRMINOS NO NEGATIVOS

La más sencilla de todas es, quizá, la serie geométrica.

[1] La expresión «no negativo» se refiere siempre a los números *reales*.

3.26 Teorema *Si* $0 \leq x < 1$, *será*

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

Si $x \geq 1$, *la serie diverge.*

Demostración Si $x \neq 1$

$$s_n = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}.$$

de donde se deduce el resultado, si hacemos $n \to \infty$. Para $x = 1$, tenemos

$$1 + 1 + 1 + \cdots,$$

que diverge, evidentemente.

En muchos casos que se encuentran en las aplicaciones, los términos de la serie decrecen monótonamente. Por tanto, es de particular interés el siguiente teorema de Cauchy. Lo que llama la atención del teorema es que una subsucesión más bien «estrecha» de $\{a_n\}$ determina la convergencia o divergencia de Σa_n.

3.27 Teorema. *Supongamos* $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$. *La serie* $\sum_{n=1}^{\infty} a_n$ *converge si, y solo si la serie*

(7) $$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \cdots$$

converge.

Demostración Por el Teorema 3.24, basta considerar el carácter de acotado de las sumas parciales. Sea

$$s_n = a_1 + a_2 + \cdots + a_n,$$
$$t_k = a_1 + 2a_2 + \cdots + 2^k a_{2^k}.$$

Para $n < 2^k$

$$\begin{aligned} s_n &\leq a_1 + (a_2 + a_3) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \\ &\leq a_1 + 2a_2 + \cdots + 2^k a_{2^k} \\ &= t_k, \end{aligned}$$

de modo que

(8) $$s_n \le t_k.$$

Por otro lado, si $n > 2^k$

$$\begin{aligned} s_n &\ge a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{k-1}+1} + \cdots + a_{2^k}) \\ &\ge \tfrac{1}{2}a_1 + a_2 + 2a_4 + \cdots + 2^{k-1}a_{2^k} \\ &= \tfrac{1}{2}t_k, \end{aligned}$$

de forma que

(9) $$2s_n \ge t_k.$$

Por (8) y(9), las sucesiones $\{s_n\}$ y $\{t_k\}$ son, o ambas acotadas o ambas no acotadas, lo que completa la demostrción.

3.28 Teorema $\sum \frac{1}{n^p}$ *converge si* $p > 1$ *y diverge si* $p \le 1$.

Demostración Si $p \le 0$, del Teorema 3.23 se deduce la divergencia. Si $p > 0$, es aplicable el Teorema 3.27 y llegamos a la serie

$$\sum_{k=0}^{\infty} 2^k \cdot \frac{1}{2^{kp}} = \sum_{k=0}^{\infty} 2^{(1-p)k}.$$

Ahora, $2^{1-p} < 1$ si, y solo si $1 - p < 0$ y se deduce la conclusión, por comparación con la serie geométrica (hágase $x = 2^{1-p}$en el Teorema 3.26).

Como nueva aplicación del Teorema 3.27 demostraremos que:

3.29 Teorema *Si* $p > 1$

(10) $$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p}$$

converge; si $p \le 1$, *la serie diverge.*

Observación «log n» expresa el logaritmo de n de base e (comparar con el Ejercicio 7, Cap. 1); el número e será definido inmediatamente (ver Definición 3.30). Comenzaremos la serie por $n = 2$, pues, $\log 1 = 0$.

Demostración La monotonía de la función logarítmica (que veremos con más detalle en el Cap. 8) implica que $\{\log n\}$ es creciente. Por tan-

to, $\{1/n \log n\}$ es decreciente y podemos aplicar el Teorema 3.27 a (10), lo que nos lleva a la serie

(11) $$\sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^k(\log 2^k)^p} = \sum_{k=1}^{\infty} \frac{1}{(k \log 2)^p} = \frac{1}{(\log 2)^p} \sum_{k=1}^{\infty} \frac{1}{k^p},$$

y del Teorema 3.28 se deduce el 3.29.

Se puede, evidentemente, continuar este proceso. Por ejemplo,

(12) $$\sum_{n=3}^{\infty} \frac{1}{n \log n \log \log n}$$

diverge, mientras que

(13) $$\sum_{n=3}^{\infty} \frac{1}{n \log n (\log \log n)^2}$$

converge.

Podemos observar que los términos de la serie (12) difieren muy poco de los de la (13). No obstante, una diverge y la otra converge. Si continuamos el proceso que nos llevó del Teorema 3.28 al 3.29 y luego a (12) y (13), encontraremos parejas de series convergentes y divergentes cuyos términos todavía difieren menos que los de (12) y (13). Se puede pensar así, que debe de haber una situación límite de cierta naturaleza, una «frontera», con todas las series convergentes a un lado, y las divergentes al otro —al menos mientras se trate de series con coeficientes monótonos—. Esta noción de «frontera» es, en realidad, completamente vaga. Queremos recalcar lo siguiente: De cualquier forma que concretemos esta noción, la idea es falsa. Los Ejercicios 11(b) y 12(b) servirán de ilustración.

No queremos profundizar más en este aspecto de la teoría de la convergencia y enviamos al lector a la «Teoría y Aplicación de las Series Infinitas» de Knopp, capítulo IX, en particular el § 41.

EL NÚMERO e

3.30 Definición $e = \sum_{n=0}^{\infty} \frac{1}{n!}$.

Donde $n! = 1 \cdot 2 \cdot 3 \cdots n$, si $n \geq 1$, y $0! = 1$

Como

$$s_n = 1 + 1 + \frac{1}{1\cdot 2} + \frac{1}{1\cdot 2\cdot 3} + \cdots + \frac{1}{1\cdot 2\cdots n}$$
$$< 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}} < 3,$$

la serie converge, y tiene sentido la definición. De hecho, la serie converge muy rápidamente y nos permite calcular e con mucha precisión.

Es interesante observar que también puede ser definido e por medio de otro proceso de límites; la demostración proporciona un buen ejemplo de operaciones con límites.

3.31 Teorema $\lim\limits_{n\to\infty}\left(1+\dfrac{1}{n}\right)^n = e.$

Demostración Sea

$$s_n = \sum_{k=0}^{n}\frac{1}{k!}, \qquad t_n = \left(1+\frac{1}{n}\right)^n.$$

Por el teorema del binomio

$$t_n = 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \cdots$$
$$+ \frac{1}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{n-1}{n}\right).$$

Por tanto, $t_n \le s_n$, de modo que

(14)
$$\limsup_{n\to\infty} t_n \le e,$$

por el Teorema 3.19. Además, si $n \ge m$,

$$t_n \ge 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \cdots + \frac{1}{m!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{m-1}{n}\right).$$

Sea $n \to \infty$, conservando m fijo. Tendremos

$$\liminf_{n\to\infty} t_n \ge 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{m!},$$

de modo que

$$s_m \leq \liminf_{n\to\infty} t_n.$$

Si se hace $m \to \infty$, tenemos finalmente

(15) $$e \leq \liminf_{n\to\infty} t_n.$$

El teorema se deduce de (14) y (15).

La rapidez con que converge la serie $\sum \frac{1}{n!}$ puede comprobarse como sigue: Si s_n tiene el mismo significado que anteriormente, tendremos

$$e - s_n = \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \cdots$$
$$< \frac{1}{(n+1)!}\left\{1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \cdots\right\} = \frac{1}{n!n}$$

de modo que

(16) $$0 < e - s_n < \frac{1}{n!n}.$$

Así, por ejemplo, s_{10} da una aproximación de e con un error menor de 10^{-7}. La desigualdad (16) tiene además interés teórico, pues nos permite demostrar la irracionalidad de e con mucha facilidad.

3.32 Teorema *e es irracional.*

Demostración Supongamos que e es racional. Será $e = p/q$, siendo p y q enteros positivos. Por (16),

(17) $$0 < q!(e - s_q) < \frac{1}{q}.$$

Por la hipótesis hecha, $q!e$ es entero. Como

$$q!s_q = q!\left(1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{q!}\right)$$

es entero, vemos que $q!1(e - s_q)$ es entero.

Como $q \geq 1$, (17) implica la existencia de un entero entre 0 y 1, con lo que hemos llegado a una contradicción.

En realidad, e no es precisamente un número algebraico. Para una demostración sencilla de esto, véase la página 25 del libro de Niven, o la 176 del de Herstein, citados en la Bibliografía.

CRITERIOS DE LA RAÍZ Y DE LA RAZÓN

3.33 Teorema (Criterio de la raíz) *Dado Σa_n, hagamos* $\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$.

Tendremos

(*a*) *si* $\alpha < 1$, Σa_n *converge;*
(*b*) *si* $\alpha > 1$, Σa_n *diverge;*
(*c*) *si* $\alpha = 1$, *el método no da información.*

Demostración Si $\alpha > 1$, podemos elegir β de modo que $\alpha < \beta < 1$ y un entero N tal que

$$\sqrt[n]{|a_n|} < \beta$$

para $n \geq N$ [por el Teorema 3.17(*b*)]. Esto es, que $n \geq N$ implica

$$|a_n| < \beta^n.$$

Como $0 < \beta < 1$, $\Sigma \beta^n$ converge. La convergencia de Σa_n puede deducirse ya por el criterio de comparación.

Si $\alpha > 1$, por el Teorema 3.17, hay una sucesión $\{n_k\}$ tal que

$$\sqrt[n_k]{|a_{n_k}|} \to \alpha.$$

Por tanto $|a_n| > 1$ para infinitos valores de n, de modo que no se cumple la condición $a_n \to 0$, necesaria para la convergencia de Σa_n (Teorema 3.23).

Para demostrar (*c*), consideremos las series

$$\Sigma \frac{1}{n}, \ \Sigma \frac{1}{n^2}.$$

Para cada una de ellas $\alpha = 1$ pero la primera diverge y la segunda converge.

3.34 Teorema (Criterio de la razón) *La serie* Σa_n

(*a*) *converge, si* $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$,

(*b*) *diverge si* $\left|\dfrac{a_{n+1}}{a_n}\right| \geq$ *para* $n \geq n_0$, *donde* n_0 *es un número entero cualquiera, dado.*

Demostración Si se cumple la condición (*a*), podemos hallar $\beta < 1$, y un entero N, tal que

$$\left|\frac{a_{n+1}}{a_n}\right| < \beta$$

para $n \geq N$. En particular,

$$\begin{aligned} |a_{N+1}| &< \beta|a_N|, \\ |a_{N+2}| &< \beta|a_{N+1}| < \beta^2|a_N|, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ |a_{N+p}| &< \beta^p|a_N|. \end{aligned}$$

esto es,

$$|a_n| < |a_N|\beta^{-N} \cdot \beta^n$$

para $n \geq N$ y se deduce (*a*) por el criterio de comparación, pues $\Sigma\beta^n$ converge.

Si $|a_{n+1}| \geq |a_n|$ para $n \geq n_0$, es fácil ver que no se cumple la condición $a_n \to 0$, y se deduce (*b*).

Nota: El saber que lím $a_{n+1}/a_n = 1$ no implica nada sobre la convergencia de Σa_n. Las series $\Sigma 1/n$ y $\Sigma 1/n^2$ demuestran esto.

3.35 Ejemplo

(*a*) Consideremos la serie

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{2^3} + \frac{1}{3^3} + \frac{1}{2^4} + \frac{1}{3^4} + \cdots,$$

para la cual

$$\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \left(\frac{2}{3}\right)^n = 0,$$

$$\liminf_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \sqrt[2n]{\frac{1}{3^n}} = \frac{1}{\sqrt{3}},$$

$$\limsup_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \sqrt[2n]{\frac{1}{2^n}} = \frac{1}{\sqrt{2}},$$

$$\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \left(\frac{3}{2}\right)^n = +\infty.$$

El criterio de la raíz indica su convergencia, el de la razón no conduce a conclusiones.

(*b*) Lo mismo es cierto para la serie

$$\frac{1}{2} + 1 + \frac{1}{8} + \frac{1}{4} + \frac{1}{32} + \frac{1}{16} + \frac{1}{128} + \frac{1}{64} + \cdots,$$

donde

$$\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} = \frac{1}{8},$$

$$\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = 2,$$

pero

$$\lim \sqrt[n]{a_n} = \tfrac{1}{2}.$$

3.36 Observaciones El criterio de la razón, frecuentemente es más fácil de aplicar que el de la raíz, pues de ordinario es más sencillo calcular cocientes que raíces n-ésimas, pero sin embargo, tiene más amplia aplicación el criterio de la raíz. Precisando más: cuando el criterio de la razón demuestra la convergencia, también el de la raíz; cuando éste no conduce a conclusiones, tampoco lo hace el de la razón. Esto es consecuencia del Teorema 3.37 y se aclara con los ejemplos anteriores.

Ninguno de los dos es interesante en caso de divergencia. Los dos la deducen del hecho de que a_n no tiende a cero cuando $n \to \infty$.

3.37 Teorema *Para toda sucesión* $\{c_n\}$ *de números positivos,*

$$\liminf_{n\to\infty} \frac{c_{n+1}}{c_n} \le \liminf_{n\to\infty} \sqrt[n]{c_n},$$

$$\limsup_{n\to\infty} \sqrt[n]{c_n} \le \limsup_{n\to\infty} \frac{c_{n+1}}{c_n}.$$

Demostración Demostraremos la segunda desigualdad; la demostración de la primera es análoga. Hagamos

$$\alpha = \limsup_{n\to\infty} \frac{c_{n+1}}{c_n}.$$

Si $\alpha = +\infty$, no hay nada que demostrar. Si α es finito, elijamos $\beta > \alpha$. Existe un entero N tal que

$$\frac{c_{n+1}}{c_n} \leq \beta$$

para $n \geq N$. En particular, para todo $p > 0$.

$$c_{N+k+1} \leq \beta c_{N+k} \qquad (k = 0, 1, \ldots, p-1).$$

Multiplicando estas desigualdades, obtenemos

$$c_{N+p} \leq \beta^p c_N,$$

o

$$c_n \leq c_N \beta^{-N} \cdot \beta^n \qquad (n \geq N).$$

De aquí

$$\sqrt[n]{c_n} \leq \sqrt[n]{c_N \beta^{-N}} \cdot \beta,$$

De modo que

(18) $$\limsup_{n\to\infty} \sqrt[n]{c_n} \leq \beta,$$

por el Teorema 3.20(b). Como (18) es cierto para todo $\beta > \alpha$, tenemos

$$\limsup_{n\to\infty} \sqrt[n]{c_n} \leq \alpha.$$

SERIES DE POTENCIAS

3.38 Definición Dada una sucesión $\{c_n\}$ de números complejos, a la serie

(19) $$\sum_{n=0}^{\infty} c_n z^n$$

se le llama *serie de potencias*. A los números c_n se les llama *coeficientes* de la serie; z es un número complejo.

En general, la serie convergerá o divergerá, según el valor de z. Más concretamente, a toda serie de potencias hay asociado un círculo, llamado

círculo de convergencia, tal que (19) converge si z está en el interior del mismo y diverge si está en el exterior (para tener en cuenta todos los casos, debemos considerar el plano como el interior de un círculo de radio infinito, y un punto como un círculo de radio cero). El comportamiento en el círculo de convergencia es mucho más variado y no se puede definir tan sencillamente.

3.39 Teorema *Dada la serie de potencias* $\Sigma c_n z^n$, *hagamos*

$$\alpha = \lim_{n\to\infty} \sup \sqrt[n]{|c_n|}, \qquad R = \frac{1}{\alpha}\cdot$$

(Si $\alpha = 0$, $R = +\infty$; *si* $\alpha = +\infty$, $R = 0$.) $\Sigma c_n z^n$ *converge si* $|z| < R$ *y diverge si* $|z| > R$.

Demostración Hagamos $a_n = c_n z^n$ y apliquemos el criterio de la raíz:

$$\lim_{n\to\infty} \sup \sqrt[n]{|a_n|} = |z| \lim_{n\to\infty} \sup \sqrt[n]{|c_n|} = \frac{|z|}{R}\cdot$$

Nota: A R se le llama radio de convergencia de $\Sigma c_n z^n$.

3.40 Ejemplos

(a) Para $\Sigma n^n z^n$: $R = 0$.

(b) Para $\sum \frac{z^n}{n!}$: $R = +\infty$ (en este caso es más sencillo de aplicar el criterio de la razón que el de la raíz).

(c) Para Σz^n: $R = 1$. Si $|z| = 1$, la serie diverge, pues $\{z^n\}$ no tiende a 0 cuando $n \to \infty$.

(d) Para $\sum \frac{z^n}{n}$: $R = 1$. En el círculo de convergencia, la serie diverge para $z = 1$, y converge para todos los demás puntos de $|z| = 1$. Esta última afirmación será demostrada en el Teorema 3.44.

(e) Para $\sum \frac{z^n}{n^2}$: $R = 1$. La serie converge también en todos los puntos del círculo $|z| = 1$, por el criterio de comparación, ya que $|z^n/n^2| = 1/n^2$ si $|z| = 1$.

SUMA POR PARTES

3.41 Teorema *Dadas dos sucesiones* $\{a_n\}$, $\{b_n\}$, *hagamos*

$$A_n = \sum_{k=0}^{n} a_k$$

si $n \geq 0$; *pongamos* $A_{-1} = 0$. *Si* $0 \leq p \leq q$, *tenemos*

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p. \tag{20}$$

Demostración

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q} (A_n - A_{n-1}) b_n = \sum_{n=p}^{q} A_n b_n - \sum_{n=p-1}^{q-1} A_n b_{n+1},$$

y la última expresión de la derecha se ve fácilmente que es igual al segundo miembro de (20).

La fórmula (20), llamada «fórmula de sumación parcial» es útil en el estudio de series de la forma $\Sigma a_n b_n$, en particular cuando $\{b_n\}$ es monótona. Veremos, ahora, algunas aplicaciones.

3.42 Teorema *Supongamos*

(*a*) *las sumas parciales* A_n de Σa_n *forman una sucesión acotada;*
(*b*) $b_0 \geq b_1 \geq b_2 \geq \cdots$;
(*c*) $\lim\limits_{n \to \infty} b_n = 0$.

Entonces, $\Sigma a_n b_n$ *converge.*

Demostración Elijamos M de modo que $|A_n| \leq M$ para todo n. Dado $\epsilon > 0$, existe un entero N tal que $b_N \leq (\varepsilon/2M)$. Para $N \leq p \leq q$, tenemos

$$\begin{aligned}\left|\sum_{n=p}^{q} a_n b_n\right| &= \left|\sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p\right| \\ &\leq M \left|\sum_{n=p}^{q-1} (b_n - b_{n+1}) + b_q + b_p\right| \\ &= 2Mb_p \leq 2Mb_N \leq \varepsilon.\end{aligned}$$

Deduciéndose la convergencia a partir del criterio de Cauchy. Observamos que la primera desigualdad de la cadena anterior depende del hecho de ser $b_n - b_{n+1} \geq 0$.

3.43 Teorema *Supongamos que*

(*a*) $|c_1| \geq |c_2| \geq |c_3| \geq \cdots$;
(*b*) $c_{2m-1} \geq 0$, $c_{2m} \leq 0$ $\quad (m = 1, 2, 3, \ldots)$;
(*c*) $\lim_{n \to \infty} c_n = 0$.

Entonces, Σc_n converge.

Las series para las que se cumple (*b*) se llaman «series alternantes»; el teorema es debido a Leibnitz.

Demostración Basta aplicar el Teorema 3.42 con $a_n = (-1)^{n+1}$, y $b_n = |c_n|$.

3.44 Teorema *Supongamos que el radio de convergencia de $\Sigma c_n z^n$ es 1, que $c_0 \geq c_1 \geq c_2 \geq \cdots$, y $\lim_{n\to\infty} c_n = 0$. $\Sigma c_n z^n$ converge en todo punto del círculo $|z| = 1$, excepto, posiblemente, en $z = 1$.*

Demostración Hagamos $a_n = z^n$, $b_n = c_n$. Se satisfacen las hipótesis del Teorema 3.42, pues

$$|A_n| = \left|\sum_{m=0}^{n} z^m\right| = \left|\frac{1 - z^{m+1}}{1 - z}\right| \leq \frac{2}{|1 - z|},$$

si $|z| = 1$, $z \neq 1$.

CONVERGENCIA ABSOLUTA

Se dice que la serie Σa_n *converge absolutamente* si converge la serie $\Sigma |a_n|$

3.45 Teorema *Si Σa_n converge absolutamente, Σa_n converge.*

Demostración La afirmación se deduce de la desigualdad

$$\left|\sum_{k=n}^{m} a_k\right| \leq \sum_{n=k}^{m} |a_k|,$$

más el criterio de Cauchy.

3.46 Observaciones Para las series de términos positivos, la convergencia absoluta coincide con la convergencia.

Si Σa_n converge, pero $\Sigma |a_n|$ diverge, decimos que Σa_n converge *no absolutamente*. Por ejemplo, la serie

$$\sum \frac{(-1)^n}{n}$$

converge no absolutamente (Teorema 3.43).

El criterio de comparación, lo mismo que el de la raíz y el de la razón, son en realidad métodos para hallar la convergencia absoluta y no pueden

dar ninguna información para las series convergentes no absolutamente. Para salvar esta dificultad se utiliza a veces la suma por partes. En particular, las series de potencias convergen absolutamente en el interior del círculo de convergencia.

Veremos que con las series absolutamente convergentes se puede operar en gran manera como con las sumas finitas: podemos multiplicarlas término a término y variar el orden en que se efectúan las sumas, sin afectar a la suma de la serie. Por el contrario, para las series no absolutamente convergentes esto ya no es verdad y hay que tomar más precauciones al operar con ellas.

ADICIÓN Y MULTIPLICACIÓN DE SERIES

3.47 Teorema *Si* $\Sigma a_n = A$, *y* $\Sigma b_n = B$, *será* $\Sigma(a_n + b_n) = A + B$, *y* $\Sigma ca_n = cA$, *para todo c prefijado.*

Demostración Sea

$$A_n = \sum_{k=0}^{n} a_k, \qquad B_n = \sum_{k=0}^{n} b_k.$$

de aquí

$$A_n + B_n = \sum_{k=0}^{n} (a_k + b_k).$$

Como $\lim_{n\to\infty} A_n = A$ y $\lim_{n\to\infty} B_n = B$, vemos que

$$\lim_{n\to\infty} (A_n + B_n) = A + B.$$

La demostración de la segunda afirmación es aún más sencilla.

Así, pues, pueden sumarse término a término dos series convergentes y la serie resultante converge hacia la suma de las dos series. El caso es más complicado cuando consideramos la multiplicación de dos series: para empezar, debemos definir el producto, lo que puede hacerse de varios modos; consideraremos el llamado «producto de Cauchy».

3.48 Definición Dadas Σa_n y Σb_n, hacemos

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} \qquad (n = 0, 1, 2, \ldots)$$

y llamamos a Σc_n, *producto* de las dos series dadas.

Se puede justificar esta definición del modo siguiente: si tenemos dos series de potencias $\Sigma a_n z^n$ y $\Sigma b_n z^n$, las multiplicaciones término a término y agrupamos las que tienen la misma potencia de z, obtenemos

$$\begin{aligned}\sum_{n=0}^{\infty} a_n z^n \cdot \sum_{n=0}^{\infty} b_n z^n &= (a_0 + a_1 z + a_2 z^2 + \cdots)(b_0 + b_1 z + b_2 z^2 + \cdots)\\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0)z + (a_0 b_2 + a_1 b_1 + a_2 b_0)z^2 + \cdots\\ &= c_0 + c_1 z + c_2 z^2 + \cdots .\end{aligned}$$

Haciendo $z = 1$, llegamos a la definición anterior.

3.49 Ejemplo Si

$$A_n = \sum_{k=0}^{n} a_k, \qquad B_n = \sum_{k=0}^{n} b_k, \qquad C_n = \sum_{k=0}^{n} c_k,$$

y $A_n \to A$ y $B_n \to B$, no es inmediato que $\{C_n\}$ converja hacia AB, pues no es $C_n = A_m B_n$. La dependencia de $\{C_n\}$ respecto a $\{A_n\}$ y $\{B_n\}$ es complicada (ver la demostración del Teorema 3.50). Vamos a ver que el producto de dos series convergentes puede, de hecho, ser divergente.

La serie

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{\sqrt{n+1}} = 1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \cdots$$

converge (Teorema 3.43). Formemos el producto de esta serie consigo misma, obteniendo

$$\begin{aligned}\sum_{n=0}^{\infty} c_n = 1 - \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\right) + \left(\frac{1}{\sqrt{3}} + \frac{1}{\sqrt{2}\sqrt{2}} + \frac{1}{\sqrt{3}}\right)&\\ - \left(\frac{1}{\sqrt{4}} + \frac{1}{\sqrt{3}\sqrt{2}} + \frac{1}{\sqrt{2}\sqrt{3}} + \frac{1}{\sqrt{4}}\right) + \cdots&,\end{aligned}$$

de forma que

$$c_n = (-1)^n \sum_{k=0}^{n} \frac{1}{\sqrt{(n-k+1)(k+1)}}.$$

Como

$$(n-k+1)(k+1) = \left(\frac{n}{2}+1\right)^2 - \left(\frac{n}{2}-k\right)^2 \le \left(\frac{n}{2}+1\right)^2.$$

tendremos

$$|c_n| \geq \sum_{k=0}^{n} \frac{2}{n+2} = \frac{2(n+1)}{n+2},$$

con lo que se ve que la condición $c_n \to 0$, necesaria para la convergencia de Σc_n, no se cumple.

Antes de ver el teorema siguiente, debido a Mertens, debemos hacer notar que hemos considerado el producto de dos series no absolutamente convergentes.

3.50 Teorema *Supongamos que*

(*a*) $\sum_{n=0}^{\infty} a_n$ *converge absolutamente,*

(*b*) $\sum_{n=0}^{\infty} a_n = A,$

(*c*) $\sum_{n=0}^{\infty} b_n = B,$

(*d*) $c_n = \sum_{k=0}^{n} a_k b_{n-k} \qquad (n = 0, 1, 2, \ldots).$

Se verificará que

$$\sum_{n=0}^{\infty} c_n = AB.$$

Esto es, el producto de dos series convergentes converge, haciéndolo además hacia el valor previsto, si al menos una de las dos series converge absolutamente.

Demostración Hagamos

$$A_n = \sum_{k=0}^{n} a_k, \qquad B_n = \sum_{k=0}^{n} b_k, \qquad C_n = \sum_{k=0}^{n} c_k, \qquad \beta_n = B_n - B.$$

Será

$$\begin{aligned} C_n &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + \cdots + (a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0) \\ &= a_0 B_n + a_1 B_{n-1} + \cdots + a_n B_0 \\ &= a_0(B + \beta_n) + a_1(B + \beta_{n-1}) + \cdots + a_n(B + \beta_0) \\ &= A_n B + a_0 \beta_n + a_1 \beta_{n-1} + \cdots + a_n \beta_0 \end{aligned}$$

Pongamos

$$\gamma_n = a_0 \beta_n + a_1\beta_{n-1} + \cdots + a_n \beta_0 .$$

Queremos demostrar que $C_n \to AB$. Como $A_nB \to AB$, basta ver que

(21) $$\lim_{n\to\infty} \gamma_n = 0.$$

Hagamos

$$\alpha = \sum_{n=0}^{\infty} |a_n| .$$

[Aquí es donde utilizaremos (*a*)]. Sea dado $\varepsilon > 0$. Por (*c*), $\beta_n \to 0$. Por tanto, podemos elegir N tal que $|\beta_n| \le \varepsilon$ para $n \ge N$, en cuyo caso

$$\begin{aligned} |\gamma_n| &\le |\beta_0 a_n + \cdots + \beta_N a_{n-N}| + |\beta_{N+1}a_{n-N-1} + \cdots + \beta_n a_0| \\ &\le |\beta_0 a_n + \cdots + \beta_N a_{n-N}| + \varepsilon\alpha. \end{aligned}$$

Conservando N fijo y haciendo $n \to \infty$, tenemos

$$\limsup_{n\to\infty} |\gamma_n| \le \varepsilon\alpha,$$

ya que $a_k \to 0$ cuando $k \to \infty$. Y, al ser arbitrario ϵ, queda demostrado (21).

Otra interrogación que se presenta es si la serie Σc_n tiene, siempre que sea convergente, la suma AB. Abel demostró que la respuesta es afirmativa:

3.51 Teorema *Si las series Σa_n, Σb_n y Σc_n convergen hacia A, B, C y $c_n = a_0b_n + \cdots + a_nb_0$, es $C = AB$.*

En este caso, no es necesario hacer ninguna hipótesis sobre la convergencia absoluta. Daremos una demostración sencilla (que se basa en la continuidad de las series potenciales) después de ver el Teorema 8.2.

REORDENAMIENTOS

3.52 Definición Sea $\{k_n\}$, $n = 1, 2, 3, \ldots$, una sucesión en la que cada entero positivo aparece una vez y solo una (esto es, $\{k_n\}$ es una función 1-1 de J en J, según la notación de la Definición 2.4). Haciendo

$$a_n' = a_{k_n} \qquad (n = 1, 2, 3, \ldots),$$

decimos que $\Sigma a_n'$ es un reordenamiento de Σa_n.

Si $\{s_n\}$ y $\{s_n'\}$ son las sucesiones de las sumas parciales de Σa_n y $\Sigma a_n'$, se ve fácilmente que, en general, las dos sucesiones constan de términos completamente diferentes. Hemos llegado, ahora, al problema de determinar en qué condiciones convergerán todas los reordenamientos de una serie convergente, y si las sumas de ellas son necesariamente iguales.

3.53 Ejemplo Consideremos la serie convergente

(22) $$1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \tfrac{1}{5} - \tfrac{1}{6} + \cdots$$

y uno de sus reordenamientos,

(23) $$1 + \tfrac{1}{3} - \tfrac{1}{2} + \tfrac{1}{5} + \tfrac{1}{7} - \tfrac{1}{4} + \tfrac{1}{9} + \tfrac{1}{11} - \tfrac{1}{6} + \cdots$$

en la que siempre a dos términos positivos sigue uno negativo. Si s es la suma de (22).

$$s < 1 - \tfrac{1}{2} + \tfrac{1}{3} = \tfrac{5}{6}.$$

Como

$$\frac{1}{4k-3} + \frac{1}{4k-1} - \frac{1}{2k} > 0$$

para $k \geq 1$, vemos que $s_3' < s_6' < s_9' < \cdots$, donde s_n' es la n-ésima suma parcial de (23). Por tanto,

$$\operatorname*{lím\,sup}_{n\to\infty} s_n' > s_3' = \tfrac{5}{6},$$

esto es, (23) no converge hacia s [dejamos al lector el comprobar que (23), sin embargo, converge].

Este ejemplo ilustra el teorema siguiente, debido a Riemann.

3.54 Teorema *Sea Σa_n una serie convergente pero no absolutamente de números reales, y sean $a \leq \beta$ dos números dados (en el sistema ampliado de números reales). Esto es,*

$$-\infty \leq \alpha \leq \beta \leq \infty.$$

Existe un reordenamiento $\Sigma a_n'$, con sumas parciales, s_n', tales que:

(24) $$\operatorname*{lím\,inf}_{n\to\infty} s_n' = \alpha, \qquad \operatorname*{lím\,sup}_{n\to\infty} s_n' = \beta.$$

Demostración Sea

$$p_n = \frac{|a_n| + a_n}{2}, \qquad q_n = \frac{|a_n| - a_n}{2} \qquad (n = 1, 2, 3, \ldots).$$

Será $p_n - q_n = a_n$; $p_n + q_n = |a_n|$; $p_n \geq 0$; $q_n \geq 0$. Las series Σp_n y Σq_n serán ambas divergentes.

Porque si fueran convergentes,

$$\Sigma(p_n + q_n) = \Sigma|a_n|$$

sería convergente, contrariamente a la hipótesis. Como

$$\sum_{n=1}^{N} a_n = \sum_{n=1}^{N} (p_n - q_n) = \sum_{n=1}^{N} p_n - \sum_{n=1}^{N} q_n,$$

la divergencia de Σp_n y la convergencia de Σq_n (o viceversa) implica la divergencia de Σa_n lo que, de nuevo, va contra la hipótesis.

Representemos, ahora, por P_1, P_2, $P_3, \ldots,$ los términos no negativos de Σa_n en el orden en que figuran y Q_1, Q_2, $Q_3, \ldots,$ los valores absolutos de los términos negativos de Σa_n en su orden.

Las series ΣP_n, ΣQ_n difieren de Σp_n, Σq_n solo en términos cero y son, por tanto, divergentes.

Construyamos las sucesiones $\{m_n\}$, $\{k_n\}$ tales que la serie

$$(25) \quad P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} + P_{m_1+1} + \cdots + P_{m_2} - Q_{k_1+1} - \cdots - Q_{k_2} + \cdots,$$

que, como se ve fácilmente, es un reordenamiento de Σa_n, satisfaga a (24).

Elijamos sucesiones con valores reales $\{\alpha_n\}$, $\{\beta_n\}$ tales que $\alpha_n \to \alpha$; $\beta_n \to \beta$; $\alpha_n < \beta_n$, $\beta_1 > 0$.

Sean m_1 k_1 los menores enteros, para los que

$$P_1 + \cdots + P_{m_1} > \beta_1,$$
$$P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} < \alpha_1;$$

m_2, k_2 los enteros menores, para los cuales

$$P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} + P_{m_1+1} + \cdots + P_{m_2} > \beta_2,$$
$$P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} + P_{m_1+1} + \cdots + P_{m_2} - Q_{k_1+1} - \cdots - Q_{k_2} < \alpha_2;$$

continuando de este modo, lo que es posible, porque ΣP_n y ΣQ_n divergen.

Si x_n, y_n representan las sumas parciales de (25), cuyos últimos términos son P_{m_n} y $-Q_{k_n}$, será

$$|x_n - \beta_n| \leq P_{m_n}, \qquad |y_n - \alpha_n| \leq Q_{k_n}.$$

Como $P_n \to 0$ y $Q_n \to 0$ cuando $n \to \infty$; vemos que $x_n \to \beta$, e $y_n \to \alpha$.

Finalmente, está claro que ningún número menor que α o mayor que β puede ser límite subsecuencial de las sumas parciales de (25).

3.55 Teorema *Si Σa_n es una serie de números complejos que converge absolutamente, entonces cualquier reordenamiento de Σa_n converge, y todos ellos convergen a la misma suma.*

Demostración Sea un reordenamiento $\Sigma a_n'$, con sumas parciales s_n'. Dado $\varepsilon > 0$, existe un entero N tal que $m \geq n \geq N$ implica

(26) $$\sum_{i=n}^{m} |a_i| \leq \varepsilon.$$

Ahora se escoge p de tal forma que los enteros $1, 2, \ldots, N$ estén todos en el conjunto $k_1, k_2, \ldots, k_p$ (se está usando la notación de la Definición 3.52). Entonces si $n > p$, los números $a_1, \ldots, a_N$ se cancelarán al llevar a cabo la diferencia $s_n - s_n'$, de tal forma que por (26), $|s_n - s_n'| \leq \varepsilon$. Por consiguiente $\{s_n'\}$ converge hacia la misma suma que $\{s_n\}$.

EJERCICIOS

1. Demostrar que la convergencia de $\{s_n\}$ implica la de $\{|s_n|\}$. ¿Es verdad la inversa?
2. Calcular $\lim_{n\to\infty} (\sqrt{n^2 + n} - n)$.
3. Si $s_1 = \sqrt{2}$, y

$$s_{n+1} = \sqrt{2 + \sqrt{s_n}} \qquad (n = 1, 2, 3, \ldots),$$

demostrar que $\{s_n\}$ converge, y que $s_n < 2$ para $n = 1, 2, 3, \ldots$

4. Hallar los límites superior e inferior de la sucesión $\{s_n\}$ definida por

$$s_1 = 0; \qquad s_{2m} = \frac{s_{2m-1}}{2}; \qquad s_{2m+1} = \frac{1}{2} + s_{2m}.$$

5. Para cada dos sucesiones reales $\{a_n\}$, $\{b_n\}$, tenemos

$$\lim_{n\to\infty} \sup (a_n + b_n) \leq \lim_{n\to\infty} \sup a_n + \lim_{n\to\infty} \sup b_n,$$

siempre que la suma de la derecha no sea de la forma $\infty - \infty$.

6. Averiguar el comportamiento (convergencia o divergencia) de Σa_n, si

(*a*) $a_n = \sqrt{n+1} - \sqrt{n}$;

(*b*) $a_n = \dfrac{\sqrt{n+1} - \sqrt{n}}{n}$;

(*c*) $a_n = (\sqrt[n]{n} - 1)^n$;

(*d*) $a_n = \dfrac{1}{1+z^n}$, para valores complejos de z.

7. Demostrar que la convergencia de Σa_n implica la de

$$\Sigma \frac{\sqrt{a_n}}{n},$$

si $a_n \geq 0$.

8. Si Σa_n converge y $\{b_n\}$ es monótona y acotada, $\Sigma a_n b_n$ converge.

9. Hallar el radio de convergencia de cada una de las series de potencias siguientes:

(*a*) $\sum n^3 z^n$,

(*b*) $\sum \dfrac{2^n}{n!} z^n$,

(*c*) $\sum \dfrac{2^n}{n^2} z^n$,

(*d*) $\sum \dfrac{n^3}{3^n} z^n$.

10. Supongamos que los coeficientes de la serie de potencias $\Sigma a_n z^n$ son enteros, y que un número infinito de ellos son distintos de cero. Demostrar que el radio de convergencia es, a lo más 1.

11. Suponer que $a_n > 0$, $s_n = a_1 + \cdots + a_n$, y Σa_n diverge.

(*a*) Demostrar que $\sum \dfrac{a_n}{1+a_n}$ diverge.

(*b*) Demostrar que

$$\frac{a_{N+1}}{s_{N+1}} + \cdots + \frac{a_{N+k}}{s_{N+k}} \geq 1 - \frac{s_N}{s_{N+k}}$$

y deducir que $\sum \dfrac{a_n}{s_n}$ diverge.

(*c*) Demostrar que

$$\frac{a_n}{s_n^2} \leq \frac{1}{s_{n-1}} - \frac{1}{s_n}$$

y deducir que $\sum \frac{a_n}{s_n^2}$ converge.

(*d*) ¿Qué se puede decir acerca de

$$\sum \frac{a_n}{1+na_n} \quad \text{and} \quad \sum \frac{a_n}{1+n^2a_n}?$$

12. Suponer que $a_n > 0$ y Σa_n converge. Hacer

$$r_n = \sum_{m=n}^{\infty} a_m.$$

(*a*) Demostrar que

$$\frac{a_m}{r_m} + \cdots + \frac{a_n}{r_n} > 1 - \frac{r_n}{r_m}$$

si $m < n$, y deducir que $\sum \frac{a_n}{r_n}$ diverge.

(*b*) Demostrar que

$$\frac{a_n}{\sqrt{r_n}} < 2(\sqrt{r_n} - \sqrt{r_{n+1}})$$

y deducir que $\sum \frac{a_n}{\sqrt{r_n}}$ converge.

13. Demostrar que el producto de Cauchy de dos series absolutamente convergentes, converge absolutamente.

14. Si $\{s_n\}$ es una sucesión compleja, se define su media aritmética σ_n como

$$\sigma_n = \frac{s_0 + s_1 + \cdots + s_n}{n+1} \qquad (n = 0, 1, 2, \ldots).$$

(*a*) Si lím $s_n = s$, demostrar que lím $\sigma_n = s$.
(*b*) Construir una sucesión $\{s_n\}$ que no converja, aunque $\sigma_n = 0$.
(*c*) ¿Qué ocurriría si $s_n > 0$ para todo n y lím sup $s_n = \infty$, aunque $\sigma_n = 0$?
(*d*) Hacer $a_n = s_n - s_{n-1}$, para $n \geq 1$. Mostrar que

$$s_n - \sigma_n = \frac{1}{n+1} \sum_{k=1}^{n} ka_k.$$

Suponer que lím $(na_n) = 0$ y que $\{\sigma_n\}$ convergen. Demostrar que $\{s_n\}$ converge. [Esto proporciona un inverso de (*a*), pero con la suposición adicional de que $na_n \to 0$.]
(*e*) Deducir la última conclusión a partir de una hipótesis más débil: Suponer $M < \infty$, $|na_n| \leq M$ para todo n, y lím $\sigma_n = \sigma$. Probar que lím $s_n = \sigma$, completando lo siguiente:

Si $m < n$, entonces

$$s_n - \sigma_n = \frac{m+1}{n-m}(\sigma_n - \sigma_m) + \frac{1}{n-m}\sum_{i=m+1}^{n}(s_n - s_i).$$

Para estas i,

$$|s_n - s_i| \leq \frac{(n-i)M}{i+1} \leq \frac{(n-m-1)M}{m+2}.$$

Con $\varepsilon > 0$ fijo y asociado a cada n el entero m que satisface la desigualdad

$$m \leq \frac{n-\varepsilon}{1+\varepsilon} < m+1.$$

Entonces $(m+1)/(n-m) \leq 1/\varepsilon$ y $|s_n - s_i| > M\varepsilon$. De aquí que

$$\limsup_{n\to\infty} |s_n - \sigma| \leq M\varepsilon.$$

Ya que ε fue arbitrario, lím $s_n = \sigma$.

15. Puede ampliarse la Definición 3.21 en el caso para el cual los a_n pertenecen a algún R^k prefijado. La convergencia absoluta se define como la convergencia de $\Sigma|\mathbf{a}_n|$. Mostrar que los Teoremas 3.22, 3.23, 3.25(*a*), 3.33, 3.34, 3.42, 3.45, 3.47 y 3.55 son verdaderos en este caso más general. (Sólo se requieren modificaciones ligeras en las demostraciones.)

16. Si α es un número positivo fijo, escoger $x_1 > \sqrt{\alpha}$, y definir $x_2, x_3, x_4, \ldots$, por medio de la fórmula de recurrencia

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{\alpha}{x_n}\right).$$

(*a*) Demostrar que $\{x_n\}$ decrece monótonamente y que lím $x_n = \sqrt{\alpha}$.

(*b*) Hacer $\varepsilon_n = x_n - \sqrt{\alpha}$, y mostrar que

$$\varepsilon_{n+1} = \frac{\varepsilon_n^2}{2x_n} < \frac{\varepsilon_n^2}{2\sqrt{\alpha}}$$

de modo que al poner $\beta = 2\sqrt{\alpha}$.

$$\varepsilon_{n+1} < \beta\left(\frac{\varepsilon_1}{\beta}\right)^{2^n} \qquad (n = 1, 2, 3, \ldots).$$

(*c*) Este es un buen algoritmo para calcular raíces cuadradas, debido a que la fórmula de recurrencia es sencilla y la convergencia en extremo rápida. Por ejemplo, si $\alpha = 3$ y $x_1 = 2$, mostrar que $\varepsilon_1/\beta < \frac{1}{10}$ y que por lo tanto

$$\varepsilon_5 < 4 \cdot 10^{-16}, \qquad \varepsilon_6 < 4 \cdot 10^{-32}.$$

17. Sea $\alpha > 1$ fijo. Si se toma $x_1 > \sqrt{\alpha}$ y se define

$$x_{n+1} = \frac{\alpha + x_n}{1 + x_n} = x_n + \frac{\alpha - x_n^2}{1 + x_n}.$$

(a) Demostrar que $x_1 > x_3 > x_5 > \cdots$.
(b) Demostrar que $x_2 < x_4 < x_6 < \cdots$.
(c) Demostrar que $\lim x_n = \sqrt{\alpha}$.
(d) Comparar la rapidez de convergencia de este proceso con la del descrito en el Ejercicio 16.

18. Reemplazar la fórmula de recurrencia del Ejercicio 16 por

$$x_{n+1} = \frac{p-1}{p} x_n + \frac{\alpha}{p} x_n^{-p+1}$$

en donde p es un entero positivo fijo. Describir el comportamiento de las sucesiones $\{x_n\}$ resultantes.

19. Asociar a cada sucesión $a = \{\alpha_n\}$ en la cual α_n es 0 ó 2, el número real

$$x(a) = \sum_{n=1}^{\infty} \frac{\alpha_n}{3^n}.$$

Demostrar que el conjunto de todos los $x(a)$ es precisamente el conjunto de Cantor que se describió en la sección 2.44.

20. Suponer que $\{p_n\}$ es una sucesión de Cauchy en un espacio métrico X, y que alguna subsucesión $\{p_{n_i}\}$ converge hacia un punto $p \in X$. Demostrar que la sucesión completa $\{p_n\}$ converge hacia p.

21. Demostrar el siguiente teorema, análogo al 3.10(b): Si $\{E_n\}$ es una sucesión de conjuntos cerrados y acotados en un espacio métrico *completo* X, $E_n \supset E_{n+1}$, y si

$$\lim_{n\to\infty} \text{diám } E_n = 0,$$

$\bigcap_1^\infty E_n$ consta solamente de un punto.

22. Supongamos que X es un espacio métrico completo, y $\{G_n\}$ una sucesión de subconjuntos abiertos densos de X. Demostrar el teorema de Baire, es decir, que $\bigcap_1^\infty G_n$ es no vacío. (De hecho, es denso en X.) *Sugerencia*: Hallar una sucesión decreciente de entornos cerrados E_n, de modo que $\overline{E}_n \subset G_n$, y aplicar el Ejercicio 21.

23. Supongamos que $\{p_n\}$ y $\{q_n\}$ son sucesiones de Cauchy en un espacio métrico X. Demostrar que la sucesión $\{d(p_n, q_n)\}$ converge. *Sugerencia*: Para cada m, n;

$$d(p_n, q_n) \leq d(p_n, p_m) + d(p_m, q_m) + d(q_m, q_n);$$

se deduce que

$$|d(p_n, q_n) - d(p_m, q_m)|$$

es pequeño si m y n son grandes.

24. Sea X un espacio métrico.

(*a*) Llamando a dos sucesiones de Cauchy $\{p_n\}$, $\{q_n\}$ *equivalentes* en X, si

$$\lim_{n\to\infty} d(p_n, q_n) = 0.$$

Demostrar que esta es una relación de equivalencia.

(*b*) Sea X^* el conjunto de todas las clases de equivalencia así obtenidas. Si $P \in X^*$, $Q \in X^*$, $\{p_n\} \in P$, $\{q_n\} \in Q$, definamos

$$\Delta(P, Q) = \lim_{n\to\infty} d(p_n, q_n);$$

Por el ejercicio 23, existe este límite. Demostrar que el número $\Delta(P,Q)$ permanece invariable si se sustituyen $\{p_n\}$ y $\{q_n\}$ por sucesiones equivalentes, y de aquí que Δ es una función distancia en X^*.

(*c*) Demostrar que el espacio métrico resultante X^* es completo.

(*d*) Para todo $p \in X$, existe una sucesión de Cauchy, cuyos términos son todos p; sea P_p el elemento de X^* que contiene esta sucesión. Demostrar que

$$\Delta(P_p, P_q) = d(p, q)$$

para todo $p, q \in X$. En otras palabras, la aplicación φ definida por $\varphi(p) = P_p$ es una isometría (es decir, una aplicación que conserva las distancias) de X en X^*.

(*e*) Demostrar que $\varphi(X)$ es denso en X^*, y que $\varphi(X) = X^*$ si X es completo. Por (*d*), podemos identificar X y $\varphi(X)$ y considerar a X sumergido en el espacio métrico completo X^*. A X^* le llamamos la completez† de X.

25. Sea X el espacio métrico cuyos puntos son los números racionales, con la métrica $d(x,y) = |x - y|$. ¿Cuál es la completez, X^*, de este espacio? (Comparar con el Ejercicio 24.)

† N. del E.: Se ha extendido el uso de la palabra *completez* como traducción del adjetivo *completness*. La Academia ha aceptado el término *compleción*, y existe tendencia a emplear este vocablo.

4
CONTINUIDAD

El concepto de función y parte de la terminología, relacionada con ella, se introdujeron en las Definiciones 2.1 y 2.2. Aunque (en capítulos posteriores) prestaremos especial interés a las funciones reales y complejas (esto es, funciones cuyos valores son números reales o complejos) también trataremos sobre funciones vectoriales (esto es, con valores en R^k) y con valores en un espacio métrico arbitrario. Los teoremas que trataremos de este modo general no resultan más sencillos si los limitamos, por ejemplo, a las funciones reales; y realmente simplifica y aclara el panorama el descartar las hipótesis innecesarias, y plantear y demostrar los teoremas en un sentido convenientemente general.

Los dominios de definición de nuestras funciones serán, pues, espacios métricos, convenientemente especificados en varios casos.

LÍMITES DE FUNCIONES

4.1 Definición Sean X y Y espacios métricos; supongamos que $E \subset X$, f mapea (o aplica) E en Y y p es un punto límite de E. Escribiremos $f(x) \to q$ cuando $x \to p$ o

(1) $$\lim_{x \to p} f(x) = q$$

si existe un punto $q \in Y$ con la siguiente propiedad: Para todo $\varepsilon > 0$ existe un $\delta > 0$, tal que

(2) $$d_Y(f(x), q) < \varepsilon$$

para todos los puntos $x \in E$, para los cuales

(3) $$0 < d_X(x, p) < \delta.$$

Los símbolos d_X y d_Y se refieren a las distancias en X e Y, respectivamente.

Si X y/o Y se sustituyen por la recta real, el plano complejo, o algún espacio euclidiano R^k, las distancias d_X, d_Y se sustituyen por los valores absolutos, o por las normas apropiadas (ver Sec. 2.16).

Debe observarse que $p \in X$, pero p no necesita ser punto de E en la definición anterior. Además, aun si $p \in E$, podemos tener $f(p) \neq \lim f(x)$.

Podemos enunciar de nuevo esta definición, en lenguaje de límites de sucesiones:

4.2 Teorema *Sean X, Y, E, f, p como en la Definición* 4.1. *Será*:

(4) $$\lim_{x \to p} f(x) = q$$

si y solo si

(5) $$\lim_{n \to \infty} f(p_n) = q$$

para toda sucesión $\{p_n\}$ en E, tal que

(6) $$p_n \neq p, \qquad \lim_{n \to \infty} p_n = p.$$

Demostración Supongamos que se cumple (4). Elijamos $\{p_n\}$ en E, de modo que satisfaga a (6). Sea $\varepsilon > 0$ dado. Existe $\delta > 0$, tal que $d_Y(f(x),q) < \varepsilon$ si $x \in E$ y $0 < d_X(x,p) < \delta$. Existe, además, un N, tal que $n > N$ implica $0 < d_X(p_n,p) < \delta$. Así, pues, para $n > N$, tenemos $d_Y(p_n),q) < \varepsilon$, lo que demuestra que se cumple (5).

Inversamente, supongamos que (4) es falso. Existirá algún $\varepsilon > 0$, tal que para todo $\delta > 0$ existe un punto $x \in E$ (dependiente de δ), para el cual $d_Y(f)x),q) \geq \varepsilon$, pero $0 \leq d_X(x,p) < \delta$. Tomando $\delta_n = 1/_n$ ($n = 1, 2, 3, \ldots$), hallamos una sucesión E que satisface a (6), para la cual (5) es falso.

Corolario *Si f tiene un límite en p, este límite es único.*

Se deduce de los Teoremas 3.2(*b*) y 4.2.

4.3 Definición Supongamos que tenemos dos funciones complejas, f y g, definidas ambas en E. Por $f + g$ significamos la función que asigna, a cada punto x de E el número $f(x) + g(x)$. De igual modo definimos la diferencia $f - g$, el producto fg y el cociente f/g de las dos funciones, debiendo entenderse que se define el cociente solamente en los puntos x de E, para los cuales $g(x) \neq 0$. Si f asigna a cada punto x de E el mismo número c, se dice que f es una función constante, o simplemente una constante, y escribimos $f = c$. Si f y g son funciones reales y $f(x) \geq g(x)$ para todo $x \in E$, escribiremos a veces $f \geq g$, por brevedad.

De igual modo, si $\mathbf{f}$ y $\mathbf{g}$ mapean (o aplican) E en R^k, definimos $\mathbf{f} + \mathbf{g}$ y $\mathbf{f} \cdot \mathbf{g}$ por

$$(\mathbf{f} + \mathbf{g})(x) = \mathbf{f}(x) + \mathbf{g}(x), \qquad (\mathbf{f} \cdot \mathbf{g})(x) = \mathbf{f}(x) \cdot \mathbf{g}(x);$$

y si λ es un número real $(\lambda \mathbf{f})(x) = \lambda \mathbf{f}(x)$.

4.4. Teorema *Supongamos que $E \subset X$ es un espacio métrico, p es un punto límite de E, f y g son funciones complejas de E, y*

$$\lim_{x \to p} f(x) = A, \qquad \lim_{x \to p} g(x) = B.$$

Será (*a*) $\lim_{x \to p} (f + g)(x) = A + B$;

(*b*) $\lim_{x \to p} (fg)(x) = AB$;

(*c*) $\lim_{x \to p} \left(\frac{f}{g}\right)(x) = \frac{A}{B}$, *si* $B \neq 0$.

Demostración Teniendo en cuenta el Teorema 4.2, estas afirmaciones se deducen inmediatamente de las propiedades análogas de las sucesiones (Teorema 3.3).

Observación Si $\mathbf{f}$ y $\mathbf{g}$ mapean E en R^k (*a*) continúa siendo cierto y (*b*) se convierte en

(*b'*) $\lim_{x \to p} (\mathbf{f} \cdot \mathbf{g})(x) = \mathbf{A} \cdot \mathbf{B}$.

(Compárese con el Teorema 3.4).

FUNCIONES CONTINUAS

4.5 Definición Supongamos que X y Y son espacios métricos, $E \subset X$, $p \in E$ y que f aplica E en Y. En estas condiciones, se dice que f es *continua en* p si para cada $\varepsilon > 0$ existe un $\delta > 0$, tal que

$$d_Y(f(x), f(p)) < \varepsilon$$

para todos los puntos $x \in E$, para los cuales $d_X(x,p) < \delta$.

Si f es continua en todo punto E, se dice que f es *continua en* E.

Se observará que f debe estar definida en el punto p para que sea continua en p. (Compárese con la observación que sigue a la Definición 4.1.)

Si p es un punto aislado de E, la definición implica que toda función f que tiene a E como dominio de definición es continua en p. Porque, si elegimos un $\varepsilon > 0$ cualquiera, podemos escoger $\delta > 0$, de modo que el único punto $x \in E$, para el cual $dx(x,p) < \delta$ sea $x = p$; entonces

$$d_Y(f(x), f(p)) = 0 < \varepsilon.$$

4.6 Teorema *En las condiciones enunciadas en la Definición* 4.5, *si admitimos además que p es un punto límite de E, f será continuo en p si, y solo si* $\lim_{x \to p} f(x) = f(p)$.

Demostración Se ve fácilmente si comparamos las Definiciones 4.1 y 4.5.

Volvamos, ahora, a la composición de funciones. Un breve resumen del teorema siguiente es que una función continua de una función continua, es continua.

4.7 Teorema *Supongamos que X, Y, Z son espacios métricos,* $E \subset X$, *f mapea E en Y, g mapea al rango de f,* $f(E)$ *en Z, y h es el mapeo de E en Z definida por*

$$h(x) = g(f(x)) \qquad (x \in E).$$

Si f es continua en todo punto $p \in E$ *y g es continua en el punto* $f(p)$, *h entonces es continua en p.*

Esta función h se denomina la *composición* o *función compuesta* de f y g. Se escribe comúnmente

$$h = g \circ f$$

Demostración Sea $\varepsilon > 0$ dado. Como g es continua en $f(p)$ existe *un* $\eta > 0$, tal que

$$d_Z(g(y), g(f(p))) < \varepsilon \text{ si } d_Y(y, f(p)) < \eta \quad \text{y} \quad y \in f(E).$$

Como f es continua en p, existe un $\delta > 0$, tal que

$$d_Y(f(x), f(p)) < n \text{ si } d_X(x, p) < \delta \quad \text{y} \quad x \in E.$$

Se deduce que

$$d_Z(h(x), h(p)) = d_Z(g(f(x)), g(f(p))) < \varepsilon$$

si $d_X(x,p) < \delta$, y $x \in E$. Así, pues, h es continua en p.

4.8 Teorema *Una aplicación (o mapeo) f de un espacio métrico X en un espacio métrico Y es continua en X si, y solo si $f^{-1}(V)$ es abierto en X para todo conjunto abierto V en Y.*

(Las imágenes inversas se definieron en la Definición 2.2.) Esta es una caracterización muy útil de la continuidad.

Demostración Supongamos que f es continuo en X y que V es un conjunto abierto en Y. Tenemos que demostrar que todo punto de $f^{-1}(V)$ es un punto interior de $f^{-1}(V)$. Para ello, sea $p \in X$ y $f(p) \in V$. Como V es abierto, existe un $\varepsilon > 0$, *tal que* $y \in V$ si $d_Y(f(p),y) < \varepsilon$, y como f es continua en p, existe $\delta > 0$, tal que $d_Y(f(x),f(p)) < \varepsilon$ si $d_X(x,p) < \delta$. Así, pues, $x \in f^{-1}(V)$ en cuanto $d_X(x,p) < \delta$.

Inversamente, supongamos que $f^{-1}(V)$ es abierto en X para todo conjunto abierto V en Y. Fijemos $p \in X$ y $\varepsilon > 0$, y sea V el conjunto de todo $y \in Y$, tal que $d_Y(y,f(p)) < \varepsilon$. V será abierto; por lo que $f^{-1}(V)$ es abierto, y, por tanto, existe un $\delta > 0$, tal que $x \in f^{-1}(V)$ en cuanto $d_X(p,x) < \delta$. Pero si $x \in f^{-1}(V)$, se verifica que $f(x) \in V$, de modo que $d_Y(f(x),f(p)) < \varepsilon$, lo que completa la demostración.

Corolario *Una aplicación f de un espacio métrico X en un espacio métrico Y es continua si y solo si $f^{-1}(C)$ es cerrado en X para cada conjunto cerrado C en Y.*

Esto se deduce del teorema, ya que un conjunto es cerrado si y solo si, su complemento es abierto, y como $f^{-1}(E^c) = [f^{-1}(E)]^c$ para cada $E \subset Y$.

Volvemos, ahora, a las funciones con valores complejos y vectoriales, y a las definidas en subconjuntos de R^k.

4.9 Teorema *Sean f y g funciones complejas continuas en un espacio métrico X. $f + g$, fg, y f/g son continuas en X.*

En el último caso, evidentemente, debemos suponer que $g(x) \neq 0$ para todo $x \in X$.

Demostración Respecto de los puntos de X no hay nada que demostrar. En los puntos límites, el enunciado se deduce de los Teoremas 4.4 y 4.6.

4.10 Teorema

(*a*) *Sean* $f_1, \ldots, f_k$ *funciones reales en un espacio métrico* X, *y sea* $\mathbf{f}$ *la aplicación de* X *en* R^k *definida por*

(7) $$\mathbf{f}(x) = (f_1(x), \ldots, f_k(x)) \qquad (x \in X);$$

$\mathbf{f}$ *es continua si, y solo si cada una de las funciones* $f_1, \ldots, f_k$ *es continua.*
(*b*) *Si* $\mathbf{f}$ *y* $\mathbf{g}$ *son aplicaciones continuas de* X *en* R^k, $\mathbf{f} + \mathbf{g}$ *y* $\mathbf{f} \cdot \mathbf{g}$ *son continuas en* X.

Las funciones $f_1, \ldots, f_k$ se llaman *componentes* de $\mathbf{f}$. Obsérvese que $\mathbf{f} + \mathbf{g}$ es una aplicación en R^k, mientras que $\mathbf{f} \cdot \mathbf{g}$ es una función *real* en X.

Demostración El apartado (*a*) se deduce de las desigualdades

$$|f_j(x) - f_j(y)| \leq |\mathbf{f}(x) - \mathbf{f}(y)| = \left\{ \sum_{i=1}^{k} |f_i(x) - f_i(y)|^2 \right\}^{\frac{1}{2}},$$

para $j = 1, \ldots, k$. El (*b*) se deduce de (*a*) y del Teorema 4.9.

4.11 Ejemplos Si $x_1, \ldots, x_k$ son las coordenadas del punto $\mathbf{x} \in R^k$, las funciones ϕ_i definidas por

(8) $$\phi_i(\mathbf{x}) = x_i \qquad (\mathbf{x} \in R^k)$$

son continuas en R^k, pues, la desigualdad

$$|\phi_i(\mathbf{x}) - \phi_i(\mathbf{y})| \leq |\mathbf{x} - \mathbf{y}|$$

demuestra que podemos tomar $\delta = \varepsilon$ en la Definición 4.5. A las funciones ϕ_i se les llama a veces *funciones coordenadas*.

Repitiendo la aplicación del Teorema 4.9, se ve que todo monomio

(9) $$x_1^{n_1} x_2^{n_2} \ldots x_k^{n_k}$$

en el que $n_1, \ldots, n_k$ son enteros no negativos, es continuo en R^k. Esto mismo es cierto para los productos de (9) por una constante, pues las constantes son, evidentemente, continuas. Se deduce que todo polinomio P, dado por

(10) $$P(\mathbf{x}) = \Sigma c_{n_1 \cdots n_k} x_1^{n_1} \ldots x_k^{n_k} \qquad (\mathbf{x} \in R^k),$$

es continuo en R^k. En él, los coeficientes $c_{n_1 \cdots n_k}$ son números complejos, $n_1, \ldots, n_k$ son enteros no negativos, y la suma en (10) tiene un número finito de términos.

Además, toda función racional en $x_1, \ldots, x_k$, esto es, todo cociente de dos polonomios de la forma (10), es continua en R^k cuando el denominador es diferente de cero.

De la desigualdad del triángulo se ve fácilmente que

$$(11) \qquad \big||\mathbf{x}| - |\mathbf{y}|\big| \leq |\mathbf{x} - \mathbf{y}| \quad (\mathbf{x}, \mathbf{y} \in R^k).$$

Por tanto, la aplicación $\mathbf{x} \to |\mathbf{x}|$ es una función real continua en R^k.

Si, ahora, $\mathbf{f}$ es una aplicación continua de un espacio métrico X en R^k, y ϕ está definida en X por $\phi(p) = |\mathbf{f}|(p)|$, se deduce por el Teorema 4.7, que ϕ es una función real continua en R^k.

4.12 Observación Hemos definido la noción de continuidad para funciones definidas en un *subconjunto* E de un espacio métrico X. Sin embargo, el complemento de E en X no juega ningún papel en esta definición (obsérvese que la situación era algo diferente para los límites de funciones). En consecuencia, podemos descartar el complemento del dominio de definición de f, lo que significa que podemos hablar solamente de aplicaciones continuas de un espacio métrico en otro, en lugar de hacerlo de aplicaciones de subconjuntos, con lo cual se simplifican el enunciado y la demostración de algunos teoremas. Ya hemos hecho uso de esto en los Teoremas 4.8 a 4.10, y continuaremos haciéndolo en las secciones de compacticidad y conexión.

CONTINUIDAD Y COMPACTICIDAD

4.13 Definición Se dice que una aplicación $\mathbf{f}$ de un conjunto E en R^k es *acotada* si existe un número real M, tal que $|\mathbf{f}(x)| \leq M$ para todo $\mathbf{x} \in E$.

4.14 Teorema *Supongamos que f es una aplicación continua de un espacio métrico compacto X en un espacio métrico Y. $f(X)$ será compacto.*

Demostración Sea $\{V_\alpha\}$ una cubierta abierta de $f(X)$. Como f es continua, el Teorema 4.8 demuestra que todos los conjuntos $f^{-1}(V_\alpha)$ son abiertos. Como X es compacto, hay un número finito de índices, $\alpha_1, \ldots, \alpha_n$, tales que

$$(12) \qquad X \subset f^{-1}(V_{\alpha_1}) \cup \cdots \cup f^{-1}(V_{\alpha_n}).$$

Como $f(f^{-1}(E)) \subset E$, para todo $E \subset Y$, (12) implica que

$$(13) \qquad f(X) \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}.$$

lo que complementa la demostración.

Nota: Hemos utilizado la expresión $f(f^{-1}(E)) \subset E$, válida para $E \subset Y$. Si $E \subset X$, solo podemos afirmar que $f^{-1}(f(E)) \supset E$ sin que, necesariamente, se cumpla la igualdad.

Ahora. deduciremos algunas consecuencias del Teorema 4.14.

4.15 Teorema *Si f es una aplicación continua de un espacio métrico compacto X en R^k,* $f(X)$ *es cerrado y acotado. Así pues,* f *es acotada.*

Se deduce del Teorema 2.41. El resultado es particularmente importante cuando f es real:

4.16 Teorema *Supongamos que f es una función real continua en un espacio métrico compacto X, y*

(14) $$M = \sup_{p \in X} f(p), \qquad m = \inf_{p \in X} f(p).$$

Existen puntos p, $q \in X$, tales que $f(p) = M$ y $f(q) = m$.

La notación utilizada en (14) significa que M es la mínima cota superior del conjunto de todos los números $f(p)$, cuando p tiene por rango a X, y m es la máxima cota inferior de este conjunto de números.

También puede enunciarse la conclusión como sigue: *Existen puntos p y q en X tales que $f(q) \leq f(x) \geq f(p)$ para todo $x \in X$*; esto es, f alcanza su máximo (en p) y su mínimo (en q).

Demostración Por el Teorema 4.15, $f(X)$ es un conjunto cerrado y acotado de números reales; por tanto, contiene a su extremo superior M y al inferior m (Teorema 2.28).

4.17 Teorema *Supongamos que f es una aplicación continua* 1-1 *de un espacio métrico compacto X sobre un espacio métrico Y. La aplicación inversa f^{-1} definida en Y por*

$$f^{-1}(f(x)) = x \qquad (x \in X)$$

es una aplicación continua de Y sobre X.

Demostración Aplicando el Teorema 4.8 a f^{-1} en lugar de f, vemos que basta demostrar que $f(V)$ es un conjunto abierto en Y para todo conjunto abierto V en X. Fijemos un tal conjunto, V.

El complementario V^c de V es cerrado en X y, por tanto, compacto (Teorema 2.35), por lo cual $f(V^c)$ es un subconjunto compacto de Y (Teorema 4.14) y es cerrado en Y (Teorema 2.34). Como f es una aplicación 1-1 y sobre, $f(V)$ es el complemento de $f(V^c)$. Por tanto, $f(V)$ es abierto.

4.18 Definición Sea f una aplicación de un espacio métrico X en un espacio métrico Y. Decimos que f es *uniformemente continua* en X si para cada $\varepsilon > 0$, existe $\delta > 0$ tal que

$$d_Y(f(p), f(q)) < \varepsilon \tag{15}$$

para todos los valores de p y q en X para los que $d_X(p,q) < \delta$.

Consideremos las diferencias entre los conceptos de continuidad y continuidad uniforme. Primeramente, la continuidad uniforme es una propiedad de una función en un conjunto, mientras que la continuidad se puede definir en un solo punto; y no tiene sentido la pregunta de si una función es uniformemente continua en un cierto punto. Segundo, si f es continua en X, es posible hallar para cada $\varepsilon > 0$ y cada punto p de X, un número $\delta > 0$ que posee la propiedad enunciada en la Definición 4.5. Este δ depende de ε *y de* p. Pero si f es uniformemente continua en X, es posible hallar, para cada $\varepsilon > 0$, *un* número $\delta > 0$ que la cumpla para *todos* los puntos p de X.

Es evidente que toda función uniformemente continua es continua. Del teorema siguiente se deduce que los dos conceptos son equivalentes en los conjuntos compactos.

4.19 Teorema *Sea f una aplicación continua de un espacio métrico compacto X en un espacio métrico Y. f es uniformemente continua en X.*

Demostración Sea $\varepsilon > 0$ dado. Como f es continua, podemos asociar a cada punto $p \in X$ un número positivo $\phi(p)$ tal que

$$q \in X, d_X(p, q) < \phi(p) \quad \text{implies} \quad d_Y(f(p), f(q)) < \frac{\varepsilon}{2}. \tag{16}$$

Sea $J(p)$ el conjunto constituido por todo $q \in X$ para el cual

$$d_X(p, q) < \tfrac{1}{2}\phi(p). \tag{17}$$

Como $p \in J(p)$, la colección de todos los conjuntos $J(p)$ es un recubrimiento abierto de X; y como X es compacto, hay un conjunto finito de puntos $p_1, \ldots, p_n$ en X tal que

$$X \subset J(p_1) \cup \cdots \cup J(p_n). \tag{18}$$

Hacemos

$$\delta = \tfrac{1}{2} \min\, [\phi(p_1), \ldots, \phi(p_n)]. \tag{19}$$

Será: $\delta > 0$. (Este es un punto en que el carácter de finito del recubrimiento, inherente a la definición de compacticidad, es esencial.

El mínimo de un conjunto finito de números positivos es positivo, mientras que el inf de un conjunto infinito de números positivos puede muy bien ser 0.)

Sean, ahora, p y q puntos de X tales que $d_X(p,q) < \delta$. Por (18) hay un entero m, $1 \leq m \leq n$, tal que $p \in J(p_m)$; por tanto,

(20) $$d_X(p, p_m) < \tfrac{1}{2}\phi(p_m),$$

y tenemos también que

$$d_X(q, p_m) \leq d_X(p, q) + d_X(p, p_m) < \delta + \tfrac{1}{2}\phi(p_m) \leq \phi(p_m).$$

Finalmente, (16) demuestra que, por consiguiente,

$$d_Y(f(p), f(q)) \leq d_Y(f(p), f(p_m)) + d_Y(f(q), f(p_m)) < \varepsilon.$$

Lo que completa la demostración.

En el Ejercicio 10 se da otra demostración.

Procederemos, ahora, a demostrar que la compacticidad es esencial en las hipótesis de los Teoremas 4.14, 4.15, 4.16 y 4.19.

4.20 Teorema *Sea E un conjunto no compacto en R^1.*

(*a*) *existe una función continua en E que no está acotada;*
(*b*) *existe una función continua y acotada en E que no tiene máximo. Si, además, E es acotado,*
(*c*) *existe una función continua en E que es no uniformemente continua.*

Demostración Supongamos primeramente que E es acotado, de modo que existe un punto límite x_0 de E, que no es punto de E. Consideremos

(21) $$f(x) = \frac{1}{x - x_0} \qquad (x \in E).$$

Es continua en E (Teorema 4.9) pero, evidentemente, no es acotada. Para ver que (21) no es uniformemente continua, sean $\varepsilon > 0$ y $\delta > 0$ arbitrarios, y elijamos un punto $x \in E$ tal que $|x - x_0| < \delta$. Tomando t suficientemente próximo a x_0, podemos hacer la diferencia $|f(t) - f(x)|$ mayor que ε, aunque $|t - x| < \delta$. Como esto es cierto para todo $\delta > 0$, f no es uniformemente continua en E.

La función g dada por

(22) $$g(x) = \frac{1}{1 + (x - x_0)^2} \qquad (x \in E)$$

es continua en E, y acotada, pues $0 < g(x) < 1$. Es claro que

$$\sup_{x \in E} g(x) = 1,$$

mientras que $g(x) < 1$ para todo $x \in E$. Así pues, g no tiene máximo en E.

Habiendo demostrado el teorema para conjuntos acotados E, supongamos que E no es acotado. Entonces $f(x) = x$ demuestra (a) mientras que

(23) $$h(x) = \frac{x^2}{1 + x^2} \qquad (x \in E)$$

establece (b), pues

$$\sup_{x \in E} h(x) = 1$$

y $h(x) < 1$ para todo $x \in E$.

La afirmación (c) sería falsa si se suprimiera la acotabilidad de la hipótesis. Porque, si E es el conjunto de todos los enteros, toda función definida en él es uniformemente continua en E. Para verlo, basta tomar $\delta < 1$ en la Definición 4.18.

Concluimos esta sección demostrando que la compacticidad es también esencial en el Teorema 4.17.

4.21 Ejemplo Sea X el intervalo semi-abierto $[0,2\pi)$ en la recta real, y $\mathbf{f}$ la aplicación de X sobre el círculo Y constituido por todos los puntos cuya distancia al origen es 1, dados por

(24) $$\mathbf{f}(t) = (\cos t, \operatorname{sen} t) \qquad (0 \leq t < 2\pi).$$

La continuidad de las funciones trigonométricas coseno y seno, tanto como sus propiedades de periodicidad, se demostrarán en el capítulo 8. Admitiéndolas, es fácil ver que $\mathbf{f}$ es una aplicación continua 1-1 de X sobre Y.

Sin embargo, la aplicación inversa (que existe, pues $\mathbf{f}$ es 1-1 y *sobre*) deja de ser continua en el punto $(1,0) = \mathbf{f}(0)$. Ciertamente, X no es compacto en este ejemplo. (Puede ser interesante observar que $\mathbf{f}^{-1}$ deja de ser continua ¡a pesar del hecho de *ser* Y compacto!)

CONTINUIDAD Y CONEXIBILIDAD

4.22 Teorema *Si f es una aplicación continua de un espacio métrico X en un espacio métrico Y, y si E es un subconjunto conexo de X, entonces $f(E)$ es conexo.*

Demostración Supóngase por el contrario, que $f(E) = A \cup B$, en donde A y B son subconjuntos separados no vacíos de Y.

Si se hace $G = E \cap f^{-1}(A)$ $H = E \cap f^{-1}(B)$.

Entonces $E = G \cup H$, y ni G ni H son vacíos.

Debido a que $A \subset \bar{A}$ (la cerradura de A), se tiene que G $f^{-1}(\bar{A})$; el último conjunto es cerrado, porque f es continua; de aquí que $\bar{G} \subset f^{-1}(\bar{A})$. Y se obtiene $f(\bar{G}) \subset \bar{A}$. Como $f(H) = B$ y $\bar{A} \cap B$ es vacío, se puede concluir que $\bar{G} \cap H$ es vacío.

El mismo argumento muestra que $\bar{G} \cap H$ es vacío. Por lo tanto G y H son separados. Esto es imposible si E es conexo.

4.23 Teorema *Sea f una función real continua en el intervalo $[a,b]$. Si $f(a) < f(b)$ y c es un número tal que $f(a) < c < f(b)$, existe un punto $x \in (a,b)$ tal que $f(x) = c$.*

Evidentemente, se mantiene un resultado similar si $f(a) > f(b)$. Hablando vulgarmente, el teorema expresa que una función real continua adopta en un intervalo todos los valores intermedios.

Demostración Por el Teorema 2.47, $[a,b]$ es conexo; por tanto, el Teorema 4.22 demuestra que $f([a,b])$ es un subconjunto conexo de R^1, y la afirmación queda demostrada si consideramos nuevamente el Teorema 2.47.

4.24 Observación A primera vista, puede parecer que el Teorema 4.23 tiene un recíproco. Esto es, puede pensarse que si para cada par de puntos $x_1 < x_2$ y para todo número c comprendido entre $f(x_1)$ y $f(x_2)$, hay un punto x en (x_1,x_2) tal que $f(x) = c$, f debe ser continua.

Que esto no es así, puede deducirse del Ejemplo 4.27(d).

DISCONTINUIDADES

Si x es un punto en el dominio de definición de la función f, en el cual ésta no es continua, decimos que f es *discontinua* en x, o que *f tiene una discontinuidad* en x. Si x está definida en un intervalo o en un segmento, se suelen distinguir dos tipos de discontinuidades. Antes de dar esta clasificación, tenemos que definir los *límites por la derecha* y *por la izquierda* de f en x, que representaremos por $f(x+)$ y $f(x-)$, respectivamente.

4.25 Definición Supongamos f definida en (a,b). Consideremos todo punto x tal que $a \leq x < b$. Escribiremos

$$f(x+) = q$$

si $f(t_n) \to q$ cuando $n \to \infty$, para todas las sucesiones $\{t_n\}$ en (x,b) tales que $t_n \to x$. Para obtener la definición de $f(x-)$, para $a < x \leq b$, nos limitamos a las sucesiones $\{t_n\}$ en (a,x).

Es claro que en cada punto x de (a,b), existe $\lim_{t \to x} f(t)$ si, y solo si

$$f(x+) = f(x-) = \lim_{t \to x} f(t).$$

4.26 Definición Supongamos f definida en (a,b). Si f es discontinua en el punto x y existe $f(x+)$ y $f(x-)$ se dice que f tiene una discontinuidad de *primera clase* o una *discontinuidad* simple, en x. De otro modo, se dice que la discontinuidad es de *segunda clase.*

Hay dos formas en que la función puede tener una discontinuidad simple: o $f(x+) \neq f(x-)$ [en cuyo caso el valor de $f(x)$ carece de interés] o $f(x+) = f(x-) \neq f(x)$.

4.27 Ejemplos

(*a*) Definamos

$$f(x) = \begin{cases} 1 & (x \text{ racional}), \\ 0 & (x \text{ irracional}). \end{cases}$$

En estas condiciones, f tiene una discontinuidad de segunda clase en cada punto x, pues no existen $f(x+)$, ni $f(x-)$.

(*b*) Definamos

$$f(x) = \begin{cases} x & (x \text{ racional}), \\ 0 & (x \text{ irracional}). \end{cases}$$

f es continua en $x = 0$ y tiene una discontinuidad de segunda clase en todos los demás puntos.

(*c*) Definamos

$$f(x) = \begin{cases} x + 2 & (-3 < x < -2), \\ -x - 2 & (-2 \leq x < 0), \\ x + 2 & (0 \leq x < 1). \end{cases}$$

f tiene una discontinuidad simple en $x = 0$ y es continua en todo otro punto de $(-3,1)$.

(*d*) Definamos

$$f(x) = \begin{cases} \operatorname{sen}\dfrac{1}{x} & (x \neq 0), \\ 0 & (x = 0). \end{cases}$$

Como no existe $f(0+)$ ni $f(0-)$, f tiene una discontinuidad de segunda clase en $x = 0$. Todavía no hemos demostrado que sen x es una función continua. Si lo suponemos de momento, el Teorema 4.7 implica que f es continua en todo punto $x \neq 0$.

FUNCIONES MONÓTONAS

Estudiaremos, ahora, las funciones que no decrecen nunca (o no crecen) en un segmento dado.

4.28 Definición Sea f un número real en (a,b). Se dice que f es *monótona creciente* en (a,b) si $a < x < y < b$ implica $f(x) \leq f(y)$. Si se invierte la última desigualdad, obtenemos la definición de función *monótona decreciente.* La clase de las funciones monótonas está constituida por las crecientes y las decrecientes.

4.29 Teorema *Sea f monótona creciente en (a,b). Existen $f(x+)$ y $f(x-)$ en todo punto x de (a,b). Más preciso:*

(25) $$\sup_{a<t<x} f(t) = f(x-) \leq f(x) \leq f(x+) = \inf_{x<t<b} f(t).$$

Además, si $a < x < y < b$, será

(26) $$f(x+) \leq f(y-).$$

Evidentemente, resultados análogos se cumplen para las funciones monótonas decrecientes.

Demostración Por hipótesis, el conjunto de números $f(t)$, donde $a < t < x$ está acotado superiormente por el número $f(x)$, y por tanto tiene una mínima cota superior que llamaremos A. Evidentemente $A \leq f(x)$. Tenemos que demostrar que $A = f(x-)$.

Sea $\varepsilon > 0$ prefijado. De la definición de A como mínima cota superior se deduce que existe $\delta > 0$ tal que $a < x - \delta < x$, y

(27) $$A - \varepsilon < f(x - \delta) \leq A.$$

Como f es monótono, tenemos

(28) $$f(x - \delta) \leq f(t) \leq A \qquad (x - \delta < t < x).$$

Combinando (27) y (28), vemos que

$$|f(t) - A| < \varepsilon \qquad (x - \delta < t < x).$$

Por tanto, $f(x-) = A$.

La segunda mitad de (25) se demuestra de igual modo.

Continuando, si $a < x < y < b$, de (25) vemos que

(29) $$f(x+) = \inf_{x<t<b} f(t) = \inf_{x<t<y} f(t).$$

La última igualdad se obtiene aplicando (25) a (a,y) en lugar de (a,b). De igual modo,

(30) $$f(y-) = \sup_{a<t<y} f(t) = \sup_{x<t<y} f(t).$$

La comparación de (29) y (30) da (26).

Corolario *Las funciones monótonas no tienen discontinuidad de segunda clase.*

Este corolario implica que toda función monótona es discontinua a lo más en un conjunto numerable de puntos. En lugar de recurrir al teorema general, cuya demostración se expone en el Ejercicio 17, damos aquí una demostración sencilla, aplicable a las funciones monótonas.

4.30 Teorema *Sea f monótona en (a,b). El conjunto de puntos de (a,b) en los que f es discontinua es a lo sumo numerable.*

Demostración Supongamos, para concretar, que f es creciente y sea E el conjunto de puntos en los que f es discontinua.

Con cada punto x de E, asociamos un número racional $r(x)$ tal que

$$f(x-) < r(x) < f(x+).$$

Como $x_1 < x_2$ implica $f(x_1+) \leq f(x_2-)$, vemos que $r(x_1) \neq r(x_2)$ si $x_1 \neq x_2$.

Hemos establecido, así una correspondencia 1-1 entre el conjunto E y un subconjunto del conjunto de números racionales. Este último, como sabemos, es numerable.

4.31 Observación Se habrá notado que las discontinuidades de una función monótona no son necesariamente aisladas. De hecho, dado un subconjunto numerable cualquiera E de (a,b) que puede incluso ser denso, pode-

mos construir una función f, monótona en (a,b) discontinua en todo punto de E, y en ningún otro punto de (a,b).

Para demostrarlo, supongamos los puntos de E ordenados en una sucesión $\{x_n\}$, $n = 1, 2, 3, \ldots$ Sea $\{c_n\}$ una sucesión de números positivos tal que Σc_n converge. Definamos

$$f(x) = \sum_{x_n < x} c_n \qquad (a < x < b). \tag{31}$$

debiendo entenderse la suma como sigue: se suman los términos con índices n para los cuales $x_n < x$. Si no hay puntos x_n a la izquierda de x, la suma es vacía, y siguiendo el convenio usual, diremos que es cero. Como (31) converge absolutamente, carece de importancia el orden en que están colocados los términos.

Dejamos al lector la comprobación de las siguientes propiedades de f:

(*a*) f es monótona creciente en (a,b);
(*b*) f es discontinua en todo punto de E; en realidad,

$$f(x_n+) - f(x_n-) = c_n.$$

(*c*) f es continua en cualquier otro punto de (a,b).

Además, no es difícil ver que $f(x-) = f(x)$ en todos los puntos de (a,b). Si una función satisface esta condición, decimos que f es *continua por la izquierda*. Si se hubiera tomado la suma en (31) para todos los índices n para los que $x_n \leq x$, hubiéramos tenido $f(x+) = f(x)$ en todo punto de (a,b), esto es, f hubiera sido *continua por la derecha*.

Pueden definirse también funciones de este tipo por otro método; como ejemplo, nos referimos al Teorema 6.16.

LÍMITES INFINITOS Y LÍMITES EN EL INFINITO

Para poder operar en el sistema ampliado de los números reales, ampliaremos el alcance de la Definición 4.1, volviéndola a enunciar en términos de vecindades.

Para todo número real x, hemos definido ya una vecindad de x como un segmento $(x - \delta, x + \delta)$.

4.32 Definición Para todo número real c, el conjunto de los números reales x tales que $x > c$, se dice que es una vecindad de $+\infty$ y se escribe $(x, +\infty)$. Del mismo modo, el conjunto $(-\infty, c)$ es una vecindad de $-\infty$.

4.33 Definición Supongamos f una función real definida en E. Decimos que

$$f(t) \to A \text{ si } t \to x,$$

donde A y x pertenecen al sistema ampliado de números reales, si por cada vecindad U de A existe una vecindad V de x tal que $V \cap E$ es no vacío, y que $f(t) \in U$ para todo $t \in V \cap E$, $t \neq x$.

Fácilmente se ve que esto coincide con la Definición 4.1 cuando A y x son reales.

También es cierto el análogo al Teorema 4.4, y la demostración no contiene nada nuevo. Lo enunciamos para constancia.

4.34 Teorema *Supongamos f y g definidas en E. Supongamos*

$$f(t) \to A, \qquad g(t) \to B \qquad \text{si } t \to x.$$

Será:

(*a*) $f(t) \to A'$ *implica* $A' = A$.
(*b*) $(f+g)(t) \to A + B$,
(*c*) $(fg)(t) \to AB$,
(*d*) $(f/g)(t) \to A/B$,

siempre que estén definidos los segundos miembros de (*b*); (*c*) *y* (*d*).

Obsérvese que no están definidos $\infty - \infty$; $0 \cdot \infty$; ∞/∞ y $A/0$ (ver la Definición 1.23).

EJERCICIOS

1. Supóngase que f es una función real definida sobre R^1 que además satisface la condición

$$\lim_{h \to 0} [f(x+h) - f(x-h)] = 0$$

para cada $x \in R^1$. ¿Lo anterior implica que f es continua?

2. Si f es un mapeo continuo de un espacio métrico X en un espacio métrico Y, demostrar que

$$f(\bar{E}) \subset \overline{f(E)}$$

para cada conjunto $E \subset X$. ($\bar{E}$ denota la cerradura de E.) Por medio de un ejemplo, mostrar que $f(\bar{E})$ puede ser un subconjunto propio de $\overline{f(E)}$.

3. Sea f una función real continua sobre un espacio métrico X. Sea $Z(f)$ (el *conjunto cero* de f) el conjunto de todos los $p \in X$ en los que $f(p) = 0$. Demostrar que $Z(f)$ es cerrado.

4. Sean f y g aplicaciones continuas de un espacio métrico X en otro espacio métrico Y, y E un subconjunto denso de X. Demostrar que $f(E)$ es denso en $f(X)$. Si además $g(p) = f(p)$ para todo $p \in E$, demuéstrese que $g(p) = f(p)$ para todo p

$\in X$. (Dicho de otro modo, una aplicación continua se determina por medio de sus valores sobre un subconjunto denso de su dominio.)

5. Si f es una función real continua definida en un conjunto cerrado $E \subset R^1$, demostrar que existen funciones reales continuas g en R^1 tales que $g(x) = f(x)$ para todo $x \in E$. (A tales funciones se les llama *extensiones continuas* de f, de E a R^1.) Demostrar que el resultado es falso si se omite la palabra «cerrado». Ampliar el resultado a las funciones con valores vectoriales. *Sugerencia*: Suponer que la gráfica de g es una recta en cada segmento de los que constituyen el complemento de E (comparar con el Ejer. 29, Cap. 2). El resultado permanece cierto si se sustituye R^1 por un espacio métrico, pero la demostración no es tan sencilla.

6. Si f está definida en E, su *gráfica* es el conjunto de los puntos $(x,f(x))$, para $x \in E$. En particular, si E es el conjunto de los números reales y f tiene valores reales, la gráfica de f es un subconjunto del plano.

Suponiendo que E es compacto, demostrar que f es continua en E si, y solo si su gráfica es compacta.

7. Si $E \subset X$ y f es una función definida en X, la *restricción* de f a E es la función g cuyo dominio de definición es E, tal que $g(p) = f(p)$ para $p \in E$. Definir f y g en R^2 por: $f(0,0) = g(0,0) = 0$, $f(x,y) = xy^2/(x^2 + y^4)$, $g(x,y) = xy^2/(x^2 + y^6)$ si $(x,y) \neq (0,0)$. Demostrar que f es acotada en R^2, que g es no acotada en cualquier vecindad de $(0,0)$ y que f no es continua en $(0,0)$; y sin embargo las restricciones de f y de g a cada recta en R^2 son continuas.

8. Sea f una función real uniformemente continua en el conjunto acotado E en R^1. Demostrar que f es acotada en E.

Demostrar que la conclusión es falsa si se omite de la hipótesis el carácter de acotado de E.

9. Mostrar que en términos de diámetros de conjuntos, el requisito de la definición de la continuidad uniforme puede volverse a enunciar de la manera siguiente: para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que diám $f(E) < \varepsilon$ para todo $E \subset X$ con diám $E < \delta$.

10. Completar los detalles de la siguiente demostración alternativa del Teorema 4.19: si f no es uniformemente continua, entonces para alguna $\varepsilon > 0$ hay sucesiones $\{p_n\}$, $\{q_n\}$ en X tales que $d_X(p_n,q_n) \to 0$, pero $d_Y(f(p_n), f(q_n)) > \varepsilon$. Usar el Teorema 2.37 para llegar a una contradicción.

11. Supóngase que f es un mapeo uniformemente continuo de un espacio métrico X en un espacio métrico Y y demuéstrese que $\{f(x_n)\}$ es una sucesión de Cauchy en Y para cada sucesión de Cauchy $\{x_n\}$ en X. Usar este resultado para dar una demostración alternativa del teorema establecido en el Ejercicio 13.

12. Una función uniformemente continua de una función uniformemente continua es uniformemente continua.

Establecer lo anterior con más precisión y demostrarlo.

13. Sea E un subconjunto denso de un espacio métrico X, y sea f una función *real*-uniformemente continua definida sobre E. Demostrar que f tiene una extensión continua de E a X (para terminología véase el Ejercicio 5). (La unicidad se dedu-

ce del Ejercicio 4.) *Sugerencia*: para cada $p \in X$ y cada entero positivo n, sea $V_n(p)$ el conjunto de todos los $q \in E$ con $d(p,q) < 1/n$. Usar el Ejercicio 9 para mostrar que la intersección de las cerraduras de los conjuntos $f(V_1(p)), f(V_2(p))$, consiste de un solo punto, por ejemplo $g(p)$, de R^1. Demostrar que la función g así definida sobre X es la extensión de f deseada.

¿Puede reemplazarse el espacio rango R^1 por R^k? ¿Por cualquier espacio métrico compacto? ¿Por cualquier espacio métrico completo? ¿Por cualquier espacio métrico?

14. Sea $I = [0,1]$ el intervalo cerrado unitario. Supóngase que f es un mapeo continuo de I en I. Demostrar que $f(x) = x$ para al menos un $x \in I$.

15. Un mapeo de X en Y se dice que es *abierto* si $f(V)$ es un conjunto abierto en Y siempre que V es un conjunto abierto en X.

Demostrar que cada mapeo abierto continuo de R^1 en R^1 es monótono.

16. Representemos por $[x]$ el mayor entero contenido en x, esto es, $[x]$ es el entero tal que $x - 1 < [x] \leq x$, y llamemos $(x) = x - [x]$ a la parte fraccionaria de x. ¿Qué discontinuidades tendrán las funciones $[x]$ y (x)?

17 Sea f una función real definida en (a,b). Demostrar que el conjunto de puntos en el que f tiene una discontinuidad simple es a lo sumo numerable. *Sugerencia*: Sea E el conjunto en el cual $f(x-) < f(x+)$. Con cada punto x de E se asocia una terna (p,q,r) de números racionales tales que

(a) $f(x-) < p < f(x+)$,

(b) $a < q < t < x$ implica $f(t) < p$,

(c) $x < t < r < b$ implica $f(t) > p$.

El conjunto formado por tales ternas es numerable. Demostrar que cada terna está asociada con, a lo sumo, un punto de E. Operar en igual forma con los otros tipos posibles de discontinuidades simples.

18. Todo número racional x puede estar escrito en la forma $x = m/n$, de donde, $n > 0$, y m y n son enteros sin ningún divisor común. Cuando $x = 0$, tomamos $n = 1$. Considerar la función f definida en R^1 por

$$f(x) = \begin{cases} 0 & (x \text{ irracional}), \\ \dfrac{1}{n} & \left(x = \dfrac{m}{n}\right). \end{cases}$$

Demostrar que f es continua en todo punto irracional, y tiene una discontinuidad simple en todo punto racional.

19. Supóngase que f es una función real con dominio R^1, y que tiene la propiedad del valor intermedio: Si $f(a) < c < f(b)$, entonces $f(x) = c$ para algún x entre a y b.

Supóngase también que, para cada racional r, el conjunto de todos los x con $f(x) = r$ es cerrado.

Demostrar que f es continua.

Sugerencia: Si $x_n \to x_0$, pero $f(x_n) > r > f(x_0)$ para algún r y todo n, entonces $f(t_n) = r$ para algún t_n entre x_0 y x_h; de aquí $t_n \to x_0$. Encontrar una contradicción. (N.J. Fine, *Amer. Math. Monthly*, vol. 73, 1966, p. 782.)

20. Si E es un subconjunto que no es vacío de un espacio métrico X, se define la distancia de $x \in X$ a E por medio de

$$\rho_E(x) = \inf_{z \in E} d(x, z).$$

(*a*) Demostrar que $\rho_E(x) = 0$ si y solo si $x \in \bar{E}$.
(*b*) Demostrar que ρ_E es una función uniformemente continua en X, mostrando primero que

$$|\rho_E(x) - \rho_E(y)| \leq d(x, y)$$

para todo $x \in X$, $y \in X$.
Sugerencia: $\rho_E(x) \leq d(x,z) \leq d(x,y) + d(y,z)$, de modo que

$$\rho_E(x) \leq d(x, y) + \rho_E(y).$$

21. Suponer que K y F son conjuntos ajenos en un espacio métrico X, K es compacto y F cerrado. Demostrar que existe $\delta > 0$ tal que $d(p,q) > \delta$ si $p \in K$ y $q \in F$. *Sugerencia*: ρ_F es una función positiva continua en K.
Demostrar que la conclusión puede ser falsa para dos conjuntos cerrados ajenos si ninguno de ellos es compacto.

22. Sean A y B conjuntos cerrados no vacíos, ajenos, en un espacio métrico X. Definir

$$f(p) = \frac{\rho_A(p)}{\rho_A(p) + \rho_B(p)} \quad (p \in X).$$

Demostrar que f es una función continua en X cuyo rango está en $[0,1]$, que $f(p)$, $= 0$ precisamente en A y $f(p) = 1$ precisamente en B. Esto establece un inverso del Ejercicio 10: todo conjunto cerrado $A \subset X$ es $Z(f)$ para alguna f real continua en X. Poniendo

$$V = f^{-1}([0, \tfrac{1}{2})), \qquad W = f^{-1}((\tfrac{1}{2}, 1]),$$

demostrar que V y W son abiertos y ajenos y que $A \subset V$ y $B \subset W$. (Así, los pares de conjuntos cerrados ajenos en un espacio métrico, pueden ser recubiertos por pares de conjuntos abiertos ajenos. Esta propiedad de los espacios métricos se llama *normalidad*.)

23. Se dice que una función de valores reales f definida en (a,b) es *convexa* si

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y)$$

siempre $a < x < b$, $a < y < b$, $0 < \lambda < 1$. Demostrar que cada función convexa es continua. Demostrar después que cada función convexa creciente de una función convexa es convexa. (Por ejemplo, si f es convexa, también lo es e^f.)
Si f es convexa en (a,b) y si $a < s < t < u < b$, mostrar que

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t}.$$

24. Aceptar que f es una función real continua definida en (a,b) y tal que

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$$

para todo $x, y \in (a,b)$. Demostrar entonces que f es convexa.

25. Si $A \subset R^k$ y $B \subset R^k$, se define $A + B$ como el conjunto de todas las sumas $\mathbf{x} + \mathbf{y}$ con $\mathbf{x} \in A$, $\mathbf{y} \in B$.

(a) Si K es compacto y C es cerrado en R^k, demostrar que $K + C$ es cerrado.

Sugerencia: Tomar $\mathbf{z} \notin K + C$, y hacer que $F = \mathbf{z} - C$ represente al conjunto de todos los $\mathbf{z} - \mathbf{y}$ con $\mathbf{y} \in C$. Entonces K y F son ajenos. Escójase δ como en el Ejercicio 21. Mostrar finalmente que la bola abierta con centro en $\mathbf{z}$ y radio δ no interseca $K + C$.

(b) Sea α un número real irracional, C_1 el conjunto de todos los enteros, y sea C_2 el conjunto de todos los $n\alpha$ con $n \in C_1$. Mostrar que C_1 y C_2 son subconjuntos cerrados de R^1 cuya suma $C_1 + C_2$ *no es* cerrada, mostrando primero que $C_1 + C_2$ es un subconjunto denso numerable de R^1.

26. Supóngase que X, Y, Z son espacios métricos, y que Y es compacto. Sea f una función que mapea X en Y, y g un mapeo uno-a-uno continuo que aplica Y en Z, y hágase $h(x) = g(f(x))$ para $x \in X$.

Demostrar que f es uniformemente continuo si h es uniformemente continuo.

Sugerencia: g^{-1} tiene un dominio compacto $g(Y)$, y $f(x) = g^{-1}(h(x))$.

Demostrar también que f es continuo si h es continuo.

Mostrar (modificando el Ejemplo 4.21, o encontrando un ejemplo diferente) que la compacticidad de Y no se puede omitir de la hipótesis, inclusive cuando X y Z son compactos.

5

DIFERENCIACIÓN

En este capítulo (excepto en la sección final) limitaremos nuestra atención a las funciones *reales* definidas en intervalos o segmentos, no solo por motivos de comodidad, sino porque cuando pasamos de las funciones reales a las de valores vectoriales, aparecen notables diferencias. La diferenciación de funciones definidas en R^k se tratará en el capítulo 9.

DERIVADA DE UNA FUNCIÓN REAL

5.1 Definición Supongamos f definida (y con valores reales) en $[a,b]$. Para cada $x \in [a,b]$, formemos el cociente

(1)
$$\phi(t) = \frac{f(t) - f(x)}{t - x} \qquad (a < t < b,\ t \neq x),$$

y definamos

(2)
$$f'(x) = \lim_{t \to x} \phi(t),$$

con la condición de que exista este límite, de acuerdo con la Definición 4.1.

Así, asociamos a la función f una función f', cuyo dominio de definición es el conjunto de los puntos x en los que existe (2); a f' se le llama *derivada* de f.

Si f' está definida en un punto x, decimos que f es *diferenciable* en x.

Si f' está definida en todos los puntos de un conjunto $E \subset [a,b]$, decimos que f es diferenciable en E.

Es posible considerar en (2) límites por la derecha y por la izquierda; lo que conduce a la definición de derivadas por la derecha y por la izquierda. En particular, en los puntos límites a y b, la derivada será, si existe una derivada por la derecha, o por la izquierda, respectivamente. Sin embargo no estudiaremos este tipo de derivadas con detalle.

Si f está definida en un segmento (a,b) y $a < x < b$, $f'(x)$ está definida por las (1) y (2) anteriores, pero en este caso no están definidas $f'(a)$ y $f'(b)$.

5.2 Teorema *Supongamos f definida en $[a,b]$. Si f es diferenciable en un punto $x \in [a,b]$, es continua en x.*

Demostración Cuando $t \to x$, por el Teorema 4.4, tenemos

$$f(t) - f(x) = \frac{f(t) - f(x)}{t - x} \cdot (t - x) \to f'(x) \cdot 0 = 0.$$

El recíproco de este teorema no es cierto. Es fácil construir funciones continuas que dejan de ser diferenciables en puntos aislados. En el capítulo 7 incluso nos encontraremos con una función que es continua en toda la recta, sin ser diferenciable en ningún punto.

5.3 Teorema *Supongamos que f y g están definidas en $[a,b]$ y son diferenciables en un punto $x \in [a,b]$. $f + g$; fg, y f/g son diferenciables en x, y*

(*a*) $(f+g)'(x) = f'(x) + g'(x)$;

(*b*) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$;

(*c*) $\left(\dfrac{f}{g}\right)'(x) = \dfrac{g(x)f'(x) - g'(x)f(x)}{g^2(x)}.$

En (*c*) se supone, evidentemente, que $g(x) \neq 0$.

Demostración (*a*) se ve claramente por el Teorema 4.4. Sea $h = fg$. Será

$$h(t) - h(x) = f(t)[g(t) - g(x)] + g(x)[f(t) - f(x)].$$

Si dividimos esta expresión por $t - x$ y observamos que $f(t) \to f(x)$ cuando $t \to x$ (Teorema 5.2), se deduce (*b*). Sea, ahora $h = f/g$. Será

$$\frac{h(t)-h(x)}{t-x}=\frac{1}{g(t)g(x)}\left[g(x)\frac{f(t)-f(x)}{t-x}-f(x)\frac{g(t)-g(x)}{t-x}\right].$$

Suponiendo que $t \to x$, y aplicando los Teoremas 4.4 y 5.2, obtenemos (c).

5.4 Ejemplos Se ve fácilmente que la derivada de una constante es cero. Si f está definida por $f(x) = x$, $f'(x) = 1$. Repitiendo la aplicación de (b) y (c) se ve que x^n es diferenciable, y que su derivada es nx^{n-1} para todo entero n (si $n < 0$ debemos limitarnos a $x \neq 0$). Así, todo polinomio es diferenciable, y lo mismo sucede con las funciones racionales, excepto en los puntos en que el denominador es cero.

El teorema siguiente se conoce como «regla de la cadena» de la diferenciación. En el capítulo 9 encontraremos versiones más generales.(*)

5.5 Teorema *Supongamos que f es continua en $[a,b]$, existe $f'(x)$ en algún punto $x \in [a,b]$, g está definida en un intervalo I que contiene el rango de f, y g es diferenciable en el punto $f(x)$. Si*

$$h(t)=g(f(t)) \qquad (a \le t \le b),$$

h es diferenciable en x, y

(3) $$h'(x)=g'(f(x))f'(x).$$

Demostración Sea $y = f(x)$. Por la definición de derivada, tenemos

(4) $$f(t)-f(x)=(t-x)[f'(x)+u(t)],$$

(5) $$g(s)-g(y)=(s-y)[g'(y)+v(s)],$$

donde $t \in [a,b]$; $s \in I$, y $u(t) \to 0$ cuando $t \to x$, $v(s) \to 0$ si $s \to y$. Sea $s = f(t)$. Aplicando primeramente (5) y a continuación (4), obtenemos

$$\begin{aligned} h(t)-h(x) &= g(f(t))-g(f(x)) \\ &= [f(t)-f(x)]\cdot[g'(y)+v(s)] \\ &= (t-x)\cdot[f'(x)+u(t)]\cdot[g'(y)+v(s)], \end{aligned}$$

o, si $t \neq x$,

(6) $$\frac{h(t)-h(x)}{t-x}=[g'(y)+v(s)]\cdot[f'(x)+u(t)].$$

(*) Se refiere a la diferenciación de funciones compuestas y es probable que sea el teorema más importante referente a las derivadas.

Suponiendo que $t \to x$, vemos que $s \to y$, por la continuidad de f, así que el segundo miembro de (6) tiende a $g'(y)f'(x)$, lo que da (3).

5.6 Ejemplos

(*a*) Sea f, una función definida por

(7) $$f(x) = \begin{cases} x \operatorname{sen} \dfrac{1}{x} & (x \neq 0), \\ 0 & (x = 0). \end{cases}$$

Admitiendo que la derivada de sen x es cos x (estudiaremos las funciones trigonométricas en el Cap. 8), podemos aplicar los Teoremas 5.3 y 5.5 cuando $x \neq 0$, y obtendremos

(8) $$f'(x) = \operatorname{sen}\frac{1}{x} - \frac{1}{x}\cos\frac{1}{x} \qquad (x \neq 0).$$

Para $x = 0$ estos teoremas no pueden aplicarse, pues, $1/x$ no está definido, y recurriremos directamente a la definición: para $t \neq 0$.

$$\frac{f(t) - f(0)}{t - 0} = \operatorname{sen}\frac{1}{t}.$$

Para $t \to 0$, esta expresión no tiende a ningún límite, de modo que $f'(0)$ no existe.

(*b*) Supongamos f definida por

(9) $$f(x) = \begin{cases} x^2 \operatorname{sen}\dfrac{1}{x} & (x \neq 0), \\ 0 & (x = 0), \end{cases}$$

Como anteriormente, obtenemos

(10) $$f'(x) = 2x \operatorname{sen}\frac{1}{x} - \cos\frac{1}{x} \qquad (x \neq 0).$$

En $x = 0$, recurrimos a la definición, y obtenemos

$$\left|\frac{f(t) - f(0)}{t - 0}\right| = \left|t \operatorname{sen}\frac{1}{t}\right| \leq |t| \qquad (t \neq 0);$$

suponiendo $t \to 0$, vemos que

(11) $$f'(0) = 0.$$

Así, pues, f es diferenciable en todos los puntos x, pero f' no es una función continua, pues, cos $(1/x)$ en (10), no tiende a un límite cuando $x \to 0$.

TEOREMAS DEL VALOR MEDIO

5.7 Definición Sea f una función real definida en un espacio métrico X. Decimos que f tiene un *máximo local* en un punto $p \in X$, si existe $\delta > 0$, tal que $f(q) \leq f(p)$ para todo $q \in X$, tal que $d(p,q) < \delta$.

Del mismo modo se definen los mínimos locales.

El próximo teorema es la base de muchas aplicaciones de la diferenciación.

5.8 Teorema *Supongamos f definida en $[a,b]$; si tiene un máximo local en un punto $x \in (a,b)$, y existe $f'(x)$, es $f'(x) = 0$.*

El enunciado análogo para los mínimos locales, es también cierto.

Demostración Elijamos δ de acuerdo con la Definición 5.7, de modo que

$$a < x - \delta < x < x + \delta < b.$$

Si $x - \delta < t < x$, será

$$\frac{f(t) - f(x)}{t - x} \geq 0.$$

Suponiendo $t \to x$, vemos que $f'(x) \geq 0$.

Si $x < t < x + \delta$, será

$$\frac{f(t) - f(x)}{t - x} \leq 0,$$

lo que demuestra que $f'(x) \leq 0$. De aquí que $f'(x) = 0$.

5.9 Teorema *Si f y g son funciones reales continuas en $[a,b]$ diferenciables en (a,b), existe un punto $x \in (a,b)$, en el cual*

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x).$$

Obsérvese que no es necesaria la diferenciabilidad en los extremos.

Demostración Hagamos

$$h(t) = [f(b) - f(a)]g(t) - [g(b) - g(a)]f(t) \qquad (a \leq t \leq b).$$

h es continua en $[a,b]$, diferenciable en (a,b) y

(12) $$h(a) = f(b)g(a) - f(a)g(b) = h(b).$$

Para demostrar el teorema, debemos ver que $h'(x) = 0$ para algún $x \in (a,b)$.

Si h es constante, se cumple para todo $x \in (a,b)$. Si $h(t) > h(a)$ para algún $t \in (a,b)$, sea x un punto de $[a,b]$ en el que h alcanza su máximo (Teorema 4.16). Por (12), $x \in (a,b)$ y el Teorema 5.8 demuestra que $h'(x) = 0$. Si $h(t) < h(a)$ para algún $t \in (a,b)$, se aplica el mismo argumento, eligiendo para x un punto de $[a,b]$ en el que h alcanza su mínimo.

A este teorema se le llama a veces teorema generalizado del valor medio; al siguiente caso especial suele llamársele «el» teorema del valor medio.

5.10 Teorema *Si f es una función real continua en $[a,b]$, que es diferenciable en (a,b), existe un punto $x \in (a,b)$, en el cual*

$$f(b) - f(a) = (b - a)f'(x).$$

Demostración Hacer $g(x) = x$ en el Teorema 5.9.

5.11 Teorema *Suponiendo f diferenciable en (a,b).*

(*a*) *Si $f'(x) \geq 0$ para todo $x \in (a,b)$, f es monótona creciente.*
(*b*) *Si $f'(x) = 0$ para todo $x \in (a,b)$, f es constante.*
(*c*) *Si $f'(x) \leq 0$ para todo $x \in (a,b)$, f es monótona decreciente.*

Demostración Todas las conclusiones se pueden extraer de la ecuación

$$f(x_2) - f(x_1) = (x_2 - x_1)f'(x),$$

que es válida para cada par de números x_1, x_2 de (a,b), para *algún* x entre x_1 y x_2.

CONTINUIDAD DE LAS DERIVADAS

Hemos visto [Ejemplo 5.6(*b*)] que una función f puede tener una derivada f', que existe en todo punto, pero es discontinua en alguno de ellos, pero no toda función es una derivada. En particular, las derivadas que existen en todos los puntos de un intervalo tienen una propiedad importante en común con las funciones continuas en un intervalo: adoptan los valores intermedios (compárese con el Teorema 4.23). El enunciado preciso es el siguiente:

5.12 Teorema *Supongamos que f es una función real diferenciable en $[a,b]$ y que $f'(a) < \lambda < f'(b)$. Existe un punto $x \in (a,b)$, tal que $f'(x) = \lambda$.*

Un resultado similar se cumple, evidentemente, si $f'(a) > f'(b)$.

Demostración Hágase $g(t) = f(t) - \lambda t$. Entonces $g'(a) < 0$, de manera que $g(t_1) < g(a)$ para algún $t_1 \in (a,b)$, y $g'(b) > 0$, así que $g(t_2) < g(b)$ para algún $t_2 \in (a,b)$. Por consiguiente, g alcanza su mínimo sobre $[a,b]$ (Teorema 4.16) en algún punto x tal que $a < x < b$. Entonces por el Teorema 5.8, $g'(x) = 0$. Es por esto que $f(x) = \lambda$.

Corolario *Si f es diferenciable en $[a,b]$, f' no puede tener en $[a,b]$ ninguna discontinuidad simple.*

Pero f' puede muy bien tener discontinuidad de segunda clase.

REGLA DE L'HOSPITAL

El siguiente teorema es útil, con frecuencia, para el cálculo de límites.

5.13 Teorema *Supongamos que f y g son reales y diferenciables en (a,b), y $g'(x) \neq 0$ para todo $x \in (a,b)$, donde $-\infty \leq a < b \leq +\infty$. Supongamos*

$$\frac{f'(x)}{g'(x)} \to A \text{ si } x \to a. \tag{13}$$

Si

$$f(x) \to 0 \quad \text{y} \quad g(x) \to 0 \text{ cuando } x \to a, \tag{14}$$

o si

$$g(x) \to +\infty \text{ cuando } x \to a, \tag{15}$$

será

$$\frac{f(x)}{g(x)} \to A \text{ cuando } x \to a. \tag{16}$$

Igualmente es cierto el enunciado análogo, si $x \to b$ o si $g(x) \to -\infty$ en (15). Hacemos notar que utilizamos el concepto de límite en el sentido amplio de la Definición 4.33.

Demostración Consideramos primeramente el caso de ser $-\infty \leq A < +\infty$. Elijamos un número real q, tal que $A < q$ y un r, tal que

$A < r < q$. Por (13), existe un punto $c \in (a,b)$, tal que $a < x < c$ implica

(17) $$\frac{f'(x)}{g'(x)} < r.$$

Si $a < x < y < c$, el Teorema 5.9 demuestra que hay un punto $t \in (x,y)$, tal que

(18) $$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r.$$

Supongamos que se cumple (14). Considerando en (18) que $x \to a$, vemos que

(19) $$\frac{f(y)}{g(y)} \leq r < q \qquad (a < y < c).$$

Supongamos, ahora, que se cumple (15). Conservando y fijo en (18), podemos elegir un punto $c_1 \in (a,y)$, tal que $g(x) > g(y)$ y $g(x) > 0$ si $a < x < c_1$. Multiplicando (18) por $[g(x) - g(y)]/g(x)$, obtenemos

(20) $$\frac{f(x)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)} \qquad (a < x < c_1).$$

Si suponemos que $x \to a$ en (20); (15) demuestra que hay un punto $c_2 \in (a,c_1)$, tal que

(21) $$\frac{f(x)}{g(x)} < q \qquad (a < x < c_2).$$

Resumiendo: (19) y (21) demuestran que para cada q, sujeto solamente a la condición $A < q$, hay un punto c_2, tal que $f(x)/g(x) < q$ si $a < x < c_2$.

Del mismo modo, si $-\infty < A \leq +\infty$, y se elige p, de modo que $p < A$,

(22) $$p < \frac{f(x)}{g(x)} \qquad (a < x < c_3),$$

y de las dos afirmaciones, se deduce (16).

DERIVADAS DE ORDEN SUPERIOR

5.14 Definición Si f tiene una derivada f' en un intervalo, y f' es a su vez diferenciable, representaremos su derivada por f'' y la llamaremos derivada segunda de f. Continuando de este modo, obtenemos funciones

$$f, f', f'', f^{(3)}, \ldots, f^{(n)},$$

cada una de las cuales es la derivada de la precedente. A $f^{(n)}$ se le llama derivada n-ésima o de orden n de f.

Para que $f^{(n)}(x)$ exista en un punto x, debe existir $f^{(n-1)}(t)$ en una vecindad de x (o en una vecindad hacia un lado, si x es un extremo del intervalo en el que está definida f) y debe ser diferenciable en x. Como ha de existir $f^{(n-1)}$ en una vecindad de x, $f^{(n-2)}$ debe ser diferenciable en esa vecindad.

TEOREMA DE TAYLOR

5.15 Teorema *Supongamos que f es una función real en $[a,b]$, n es un entero positivo, $f^{(n-1)}$* es *continua en $[a,b]$, existe $f^{(n)}(t)$ para todo $t \in (a,b)$. Sean α, β puntos distintos de $[a,b]$ y definamos*

(23) $$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t-\alpha)^k.$$

Existe un punto x entre α y *β, tal que*

(24) $$f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta-\alpha)^n.$$

Para $n = 1$ este teorema es el del valor medio. En general, el teorema demuestra que se puede hallar f aproximada por medio de un polinomio de grado $n - 1$; y (24) nos permite estimar el error, si conocemos las cotas en $|f^{(n)}(x)|$.

Demostración Sea M el número definido por

(25) $$f(\beta) = P(\beta) + M(\beta-\alpha)^n$$

y hagamos

(26) $$g(t) = f(t) - P(t) - M(t-\alpha)^n \qquad (a \leq t \leq b).$$

Tenemos que demostrar que $n!M = f^{(n)}(x)$ para algún x entre α y β. Por (23) y (26)

(27) $$g^{(n)}(t) = f^{(n)}(t) - n!M \qquad (a < t < b).$$

Por tanto, la demostración estará completa si podemos probar que $g^{(n)}(x) = 0$ para algún x entre α y β.

Como $P^{(k)}(\alpha) = f^{(k)}(\alpha)$ para $k = 0, \ldots, n-1$, tenemos

(28) $$g(\alpha) = g'(\alpha) = \cdots = g^{(n-1)}(\alpha) = 0.$$

La elección de M demuestra que $g(\beta) = 0$, de modo que $g'(x_1) = 0$ para algún x_1 entre α y β, por el teorema del valor medio. Como $g'(\alpha) = 0$, deducimos del mismo modo que $g''(x_2) = 0$ para algún x_2 entre α y x_1. Después de n pasos, llegamos a la conclusión de que $g^{(n)}(x_n) = 0$ para algún x_n entre α y x_{n-1}, esto es, entre α y β.

DIFERENCIACIÓN DE FUNCIONES VECTORIALES

5.16 Observación La Definición 5.1 se aplica sin ningún cambio a las funciones complejas f definidas en $[a,b]$ y los Teoremas 5.2 y 5.3, lo mismo que sus demostraciones permanecen válidos. Si f_1 y f_2 son las partes real e imaginaria de f, esto es si $f(t) = f_1(t) + if_2(t)$ para $a \le t \le b$, donde $f_1(t)$ y $f_2(t)$ son reales, se ve fácilmente que

(29) $$f'(x) = f_1'(x) + if_2'(x);$$

además, f es diferenciable en x si, y solo si f_1 y f_2 son diferenciables en x.

Pasando a las funciones con valores vectoriales (o vectoriales simplemente) en general, esto es, a las funciones $\mathbf{f}$ que mapean $[a,b]$ en algún R^k, también se puede aplicar la Definición 5.1 para definir $\mathbf{f}'(x)$. El término $\phi(t)$ de (1) es ahora, para cada t, un punto en R^k y en (2) se toma el límite respecto a la norma de R^k. En otras palabras, $\mathbf{f}'(x)$ es el punto de R^k (si existe), para el cual

(30) $$\lim_{t \to x} \left| \frac{\mathbf{f}(t) - \mathbf{f}(x)}{t - x} - \mathbf{f}'(x) \right| = 0,$$

y $\mathbf{f}'$ es también una función con valores en R^k.

Si $f_1, \ldots, f_k$ son las componentes de $\mathbf{f}$, que se definieron en el Teorema 4.10, entonces

(31) $$\mathbf{f}' = (f_1', \ldots, f_k'),$$

y $\mathbf{f}$ es diferenciable en un punto x si, y solo si cada una de las funciones $f_1, \ldots, f_k$ es diferenciable en x.

El Teorema 5.2 es cierto, lo mismo que los 5.3(a) y (b), si se sustituye fg por el producto interno $\mathbf{f} \cdot \mathbf{g}$ (ver la Definición 4.3).

Sin embargo, respecto al teorema del valor medio y a una de sus consecuencias: la regla de L'Hospital, la situación cambia. Los dos siguientes

ejemplos demostrarán que dejan de ser ciertos para las funciones con valores complejos.

5.17 Ejemplo Definamos, para x real

(32) $$f(x) = e^{ix} = \cos x + i \operatorname{sen} x.$$

(La última expresión puede considerarse como definición de la exponencial compleja e^{ix}; véase en el Capítulo 8 un estudio completo de estas funciones). Será

(33) $$f(2\pi) - f(0) = 1 - 1 = 0,$$

pero

(34) $$f'(x) = ie^{ix},$$

de modo que $|f'(x)| = 1$ para todo x real.

Así, pues, el Teorema 5.10 deja de ser cierto en este caso.

5.18 Ejemplo En el segmento (0,1) definamos $f(x) = x$, y

(35) $$g(x) = x + x^2 e^{i/x^2}.$$

Como $|e^{it}| = 1$ para todo t real, vemos que

(36) $$\lim_{x \to 0} \frac{f(x)}{g(x)} = 1.$$

Ahora bien,

(37) $$g'(x) = 1 + \left\{2x - \frac{2i}{x}\right\} e^{i/x^2} \qquad (0 < x < 1),$$

de modo que

(38) $$|g'(x)| \geq \left|2x - \frac{2i}{x}\right| - 1 \geq \frac{2}{x} - 1.$$

De aquí que

(39) $$\left|\frac{f'(x)}{g'(x)}\right| = \frac{1}{|g'(x)|} \leq \frac{x}{2 - x}$$

y así

$$\lim_{x\to 0}\frac{f'(x)}{g'(x)} = 0. \tag{40}$$

Por (36) y (40), no se cumple en este caso la regla de L'Hospital. Obsérvese también que $g'(x) \neq 0$ en $(0,1)$, por (38).

Sin embargo, *hay* una consecuencia del teorema del valor medio, que para las aplicaciones es casi tan útil como el Teorema 5.10, y que sigue siendo verdad para las funciones vectoriales. Del Teorema 5.10 se deduce que

$$|f(b) - f(a)| \leq (b-a) \sup_{a<x<b} |f'(x)|. \tag{41}$$

5.19 Teorema *Supóngase que* $\mathbf{f}$ *es una aplicación continua de* $[a,b]$ *en* R^k *y* $\mathbf{f}$ *es diferenciable en* (a,b). *Entonces existe* $x \in (a,b)$ *tal que*

$$|\mathbf{f}(b) - \mathbf{f}(a)| \leq (b-a)|\mathbf{f}'(x)|.$$

Demostración[1] Haciendo $\mathbf{z} = \mathbf{f}(b) - \mathbf{f}(a)$, y definiendo

$$\varphi(t) = \mathbf{z}\cdot\mathbf{f}(t) \qquad (a \leq t \leq b).$$

Entonces φ es una función continua de valores reales sobre $[a,b]$ que es diferenciable en (a,b). El teorema del valor medio muestra por lo tanto que

$$\varphi(b) - \varphi(a) = (b-a)\varphi'(x) = (b-a)\mathbf{z}\cdot\mathbf{f}'(x)$$

para algún $x \in (a,b)$. Por otro lado,

$$\varphi(b) - \varphi(a) = \mathbf{z}\cdot\mathbf{f}(b) - \mathbf{z}\cdot\mathbf{f}(a) = \mathbf{z}\cdot\mathbf{z} = |\mathbf{z}|^2.$$

La desigualdad de Schwarz produce ahora

$$|\mathbf{z}|^2 = (b-a)|\mathbf{z}\cdot\mathbf{f}'(x)| \leq (b-a)|\mathbf{z}|\,|\mathbf{f}'(x)|.$$

Por consiguiente $|\mathbf{z}| \leq (b-a)|\mathbf{f}'(x)|$, que es la conclusión que se quería.

EJERCICIOS

1. Sea f definida para todo x real, y supóngase que

$$|f(x) - f(y)| \leq (x-y)^2$$

[1] V.P. Havin hizo la traducción al ruso de la segunda edición de este libro y agregó esta demostración al original.

para todo x y y reales. Demostrar que f es constante.

2. Supóngase que $f'(x) > 0$ en (a,b). Demostrar que f es estrictamente creciente en (a,b), y sea g su función inversa. Demostrar que g es diferenciable, y además que

$$g'(f(x)) = \frac{1}{f'(x)} \qquad (a < x < b).$$

3. Supóngase que g es una función real sobre R^1, con derivada acotada (es decir $|g'| \leq M$). Fíjese $\varepsilon > 0$, y defínase $f(x) = x + \varepsilon g(x)$. Demostrar que f es uno-a-uno si ε es suficientemente pequeño.
(Puede determinarse un conjunto de valores admisibles de ε que depende sólo de M.)

4. Si

$$C_0 + \frac{C_1}{2} + \cdots + \frac{C_{n-1}}{n} + \frac{C_n}{n+1} = 0,$$

donde $C_0, \ldots, C_n$ son constantes reales, demostrar que la ecuación

$$C_0 + C_1 x + \cdots + C_{n-1} x^{n-1} + C_n x^n = 0$$

tiene, al menos, una raíz real entre 0 y 1.

5. Suponer f definida y diferenciable para todo $x > 0$, y $f'(x) \to 0$ cuando $x \to + + \infty$. Hacer $g(x) = f(x + 1) - f(x)$. Demostrar que $g(x) \to 0$ cuando $x \to + \infty$.

6. Suponer que
(*a*) f es continua para $x \geq 0$,
(*b*) $f'(x)$ existe para $x > 0$,
(*c*) $f(0) = 0$,
(*d*) f' es monótona creciente.
Hacer

$$g(x) = \frac{f(x)}{x} \qquad (x > 0)$$

y demostrar que g es monótona creciente.

7. Suponer que existen $f'(x)$ y $g'(x)$, $g'(x) \neq 0$ y $f(x) = g(x) = 0$. Demostrar que

$$\lim_{t \to x} \frac{f(t)}{g(t)} = \frac{f'(x)}{g'(x)}.$$

(Esto se cumple también para funciones complejas.)

8. Suponer que f' es continua en $[a,b]$ y $\varepsilon > 0$. Demostrar que existe $\delta > 0$ tal que

$$\left| \frac{f(t) - f(x)}{t - x} - f'(x) \right| < \varepsilon$$

siempre que $0 < |t - x| < \delta$; $a \leq x \leq b$; $a \leq t \leq b$. (Podría expresarse esto diciendo que f es *uniformemente diferenciable* en $[a,b]$ si f' es continua en $[a,b]$). ¿Se cumple esto también para las funciones vectoriales?

9. Sea f una función real continua sobre R^1, de la cual se sabe que $f'(x)$ existe para todo $x \neq 0$ y que $f'(x) \to 3$ cuando $x \to 0$. ¿Puede deducirse que $f'(0)$ existe?

10. Supongamos que f y g son funciones complejas diferenciales en $(0,1)$; $f(x) \to 0$; $g(x) \to 0$; $f'(x) \to A$; $g'(x) \to B$ cuando $x \to 0$, siendo A y B números complejos y $B \neq 0$. Demostrar que

$$\lim_{x \to 0} \frac{f(x)}{g(x)} = \frac{A}{B}.$$

Comparar con el Ejercicio 5.18. *Sugerencia*:

$$\frac{f(x)}{g(x)} = \left\{\frac{f(x)}{x} - A\right\} \cdot \frac{x}{g(x)} + A \cdot \frac{x}{g(x)}.$$

Aplicar el Teorema 5.13 a las partes real e imaginaria de $f(x)/x$ y $g(x)/x$.

11. Suponer que f está definida en una vecindad de x y que existe $f''(x)$. Demostrar que

$$\lim_{h \to 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} = f''(x).$$

Demostrar con un ejemplo que puede existir el límite aún cuando no exista $f''(x)$.
Sugerencia: Aplicar el Teorema 5.13.

12. Si $f(x) = |x|^3$, calcular $f'(x)$ y $f''(x)$ para todo número real x, y demostrar que $f^{(3)}(0)$ no existe.

13. Supóngase que a y c son números reales, $c > 0$, y f está definida sobre $[-1,1]$ por medio de

$$f(x) = \begin{cases} x^a \operatorname{sen}(x^{-c}) & (\text{si } x \neq 0), \\ 0 & (\text{si } x = 0). \end{cases}$$

Demostrar lo siguiente:

(*a*) f es continua si, y solo si $a > 0$.
(*b*) $f'(0)$ existe si, y solo si $a > 1$.
(*c*) f' es acotada si, y solo si $a \geq 1 + c$.
(*d*) f' es continua si, y solo si $a > 1 + c$.
(*e*) $f''(0)$ existe si, y solo si $a > 2 + c$.
(*f*) f'' es acotada si, y solo si $a \geq 2 + 2c$.
(*g*) f'' es continua si, y solo si $a > 2 + 2c$.

14. Sea f una función real diferenciable definida en (a,b). Demostrar que f es convexa si, y solo si f' es monótona creciente. Supóngase enseguida que $f''(x)$ existe para cada $x \in (a,b)$, y demuestre que f es convexa si, y solo si $f''(x) \geq 0$ para todo $x \in (a,b)$.

15. Supóngase que $a \in R^1$, y que f es una función real diferenciable dos veces sobre (a, ∞), y M_0, M_1, M_2 son las mínimas cotas superiores de $|f(x)|$, $|f'(x)|$, $|f''(x)|$, respectivamente, sobre (a, ∞). Demostrar que

$$M_1^2 \leq 4 M_0 M_2.$$

Sugerencia: Si $h > 0$, el teorema de Taylor muestra que

$$f'(x) = \frac{1}{2h}[f(x+2h) - f(x)] - hf''(\xi)$$

para algún $\xi \in (x, x + 2h)$. En consecuencia

$$|f'(x)| \leq hM_2 + \frac{M_0}{h}.$$

Para mostrar que $M_1^2 = 4M_0M_2$ puede ocurrir en realidad, tómese $a = -1$, y defínase

$$f(x) = \begin{cases} 2x^2 - 1 & (-1 < x < 0), \\ \dfrac{x^2 - 1}{x^2 + 1} & (0 \leq x < \infty), \end{cases}$$

y muéstrese que $M_0 = 1$, $M_1 = 4$, $M_2 = 4$.

¿Es también $M_1^2 \leq 4\, M_0M_2$ válido para funciones vectoriales?

16. Supóngase que f es doblemente diferenciable sobre $(0, \infty)$, f'' es acotada sobre $(0, \infty)$, y que $f(x) \to 0$ cuando $x \to \infty$. Demostrar que $f'(x) \to 0$ cuando $x \to \infty$.

Sugerencia: Hacer $a \to \infty$ en el Ejercicio 15.

17. Supóngase que f es una función real y tres veces diferenciable sobre $[-1, 1]$, tal que

$$f(-1) = 0, \qquad f(0) = 0, \qquad f(1) = 1, \qquad f'(0) = 0.$$

Demostrar que $f^{(3)}(x) \geq 3$ para algún $x \in (-1, 1)$.

Nótese que la igualdad es válida para $\frac{1}{2}(x^3 + x^2)$.

Sugerencia: Usar el Teorema 5.15, con $\alpha = 0$ y $\beta = \pm 1$, para mostrar que existen $s \in (0, 1)$ y $t \in (-1, 0)$ tales que

$$f^{(3)}(s) + f^{(3)}(t) = 6.$$

18. Supóngase que f es una función real sobre $[a,b]$, n es un entero positivo, y $f^{(n-1)}$ existe para cada $t \in [a,b]$. Sean α, β, y P del teorema de Taylor (5.15). Defínase

$$Q(t) = \frac{f(t) - f(\beta)}{t - \beta}$$

para $t \in [a,b]$, $t \neq \beta$, diferencíese

$$f(t) - f(\beta) = (t - \beta)Q(t)$$

$n - 1$ veces en $t = \alpha$, y dedúzcase la siguiente versión del teorema de Taylor:

$$f(\beta) = P(\beta) + \frac{Q^{(n-1)}(\alpha)}{(n-1)!}(\beta - \alpha)^n.$$

19. Supóngase que f está definida en $(-1, 1)$ y $f'(0)$ existe. Supóngase también que $-1 < \alpha_n < \beta_n < 1$, $\alpha_n \to 0$, y $\beta_n \to 0$ cuando $n \to \infty$. Defínanse los cocientes de las diferencias siguientes:

$$D_n = \frac{f(\beta_n) - f(\alpha_n)}{\beta_n - \alpha_n}.$$

Demostrar que:

(a) Si $\alpha_n < 0 < \beta_n$, entonces lím $D_n = f'(0)$.

(b) Si $0 < \alpha_n < \beta_n$, y $\{\beta_n/(\beta_n - \alpha_n)\}$ es acotada, entonces lím $D_n = f'(0)$.

(c) Si f'' es continua en $(-1, 1)$, entonces lím $D_n = f'(0)$.

Dar un ejemplo en el que f sea diferenciable en $(-1, 1)$ (pero f' no séa continua en 0) y en el que α_n, β_n tienda a 0 de tal forma que lím D_n exista, pero sea diferente de $f'(0)$.

20. Formular y demostrar una desigualdad que se deduce del teorema de Taylor, y que permanece válida para las funciones vectoriales.

21. Sea E un subconjunto cerrado de R^1. Dijimos en el Ejercicio 22, capítulo 4 que existe una función real continua f en R^1 cuyo conjunto cero es E. ¿Es posible, para cada conjunto cerrado E, hallar una tal f que sea diferenciable en R^1, o una que sea n veces diferenciable o incluso una que tenga derivada de todos los órdenes en R^1?

22. Supóngase que f es una función real definida sobre $(-\infty, \infty)$. Se dirá que x es un *punto fijo* de f si $f(x) = x$.

(a) Si f es diferenciable y $f'(t) \neq 1$ para cada real t, demostrar que f tiene a lo más un punto fijo.

(b) Mostrar que la función f definida por

$$f(t) = t + (1 + e^t)^{-1}$$

no tiene punto fijo, aunque $0 < f'(t) < 1$ para todo real t.

(c) No obstante, si hay una constante $A < 1$ tal que $|f'(t)| \leq A$ para todo real t, demostrar que existe un punto fijo x de f, y que $x = \lim x_n$, en donde x_1 es un número real arbitrario y además

$$x_{n+1} = f(x_n)$$

para $n = 1, 2, 3, \ldots$.

(d) Mostrar que el procedimiento descrito en el apartado (c) puede entenderse a partir de la trayectoria en zig-zag

$$(x_1, x_2) \to (x_2, x_2) \to (x_2, x_3) \to (x_3, x_3) \to (x_3, x_4) \to \cdots .$$

23. La función f definida por

$$f(x) = \frac{x^3 + 1}{3}$$

tiene tres puntos fijos, a saber α, β, γ, donde

$$-2 < \alpha < -1, \qquad 0 < \beta < 1, \qquad 1 < \gamma < 2.$$

Si se escoge arbitrariamente x_1, y se define $\{x_n\}$ haciendo $x_{n+1} = f(x_n)$.

(a) Si $x_1 < \alpha$, demostrar que $x_n \to -\infty$ cuando $n \to \infty$.

(b) Si $\alpha < x_1 < \gamma$, demostrar que $x_n \to \beta$ cuando $n \to \infty$.

(c) Su $\gamma < x_1$, demostrar que $x_n \to +\infty$ cuando $n \to \infty$.

Entonces β puede localizarse con este método, pero α y γ no.

24. El procedimiento que se describió en la parte (c) del Ejercicio 22 puede también aplicarse a funciones que mapean $(0, \infty)$ en $(0, \infty)$.

Fijando algún $\alpha > 1$, y haciendo

$$f(x) = \frac{1}{2}\left(x + \frac{\alpha}{x}\right), \qquad g(x) = \frac{\alpha + x}{1 + x}.$$

Entonces f y g tienen como único punto fijo en $(0, \infty)$ a $\sqrt{\alpha}$. Intentar explicar, en base a las propiedades de f y g, por qué la convergencia en el Ejercicio 16 del capítulo 3 es mucho más rápida que la del Ejercicio 17. (Compárense f' y g', y dibújense las trayectorias en zig-zag que se sugieren en el Ejercicio 22.)

Hacer lo mismo cuando $0 < \alpha < 1$.

25. Supóngase que f es doblemente diferenciable sobre $[a,b]$, $f(a) < 0$, $f(b) > 0$, $f'(x) \geq \delta > 0$, y $0 \leq f''(x) \leq M$ para todo $x \in [a,b]$. Sea ξ el único punto en (a,b) en el cual $f(\xi) = 0$. Completar los detalles del siguiente esbozo del *método de Newton* para calcular ξ.

(a) Escoger $x_1 \in (\xi, b)$, y definir $\{x_n\}$ por medio de

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Interpretar esto geométricamente, en términos de una tangente a la gráfica de f.

(b) Demostrar que $x_{n+1} < x_n$ y que

$$\lim_{n \to \infty} x_n = \xi.$$

(c) Usar el teorema de Taylor para mostrar que

$$x_{n+1} - \xi = \frac{f''(t_n)}{2f'(x_n)}(x_n - \xi)^2$$

para algún $t_n \in (\xi, x_n)$.

(d) Si $A = M/2\delta$, dedúzcase que

$$0 \leq x_{n+1} - \xi \leq \frac{1}{A}[A(x_1 - \xi)]^{2^n}.$$

(Compárese con los Ejercicios 16 y 18 del capítulo 3.)

(e) Mostrar que el método de Newton es significativo para encontrar un punto fijo de la función g definida por

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

¿Cómo se comporta $g'(x)$ cuando x está muy cerca de ξ?

(f) Hacer $f(x) = x^{1/3}$ sobre $(-\infty, \infty)$ y aplicar el método de Newton. ¿Qué ocurre?

26. Suponer que f es diferenciable en $[a,b]$; $f(a) = 0$ y hay un número real A tal que

$|f'(x)| \leq A|f(x)|$ en $[a,b]$. Demostrar que $f(x) = 0$ para todo $x \in [a,b]$. *Sugerencia*: Dado $x_0 \in [a,b]$, sea $M_0 = \sup |f(x)|$,

$$M_1 = \sup|f'(x)|$$

para $a \leq x \leq x_0$. Para cada tal x,

$$|f(x)| \leq M_1(x_0 - a) \leq A(x_0 - a)M_0 .$$

Por tanto $M_0 = 0$, si $A(x_0 - a) < 1$. Esto es, $f = 0$ en $[a, x_0]$. Continuar.

27. Sea ϕ una función real definida en un rectángulo R en el plano, dado por $a \leq x \leq b$, $\alpha \leq y \leq \beta$. Una *solución* del problema con valores iniciales

$$y' = \phi(x, y), \qquad y(a) = c \qquad (\alpha \leq c \leq \beta)$$

es, por definición, una función diferenciable f en $[a,b]$ tal que $f(a) = c$, $\alpha \leq f(x) \leq \beta$, y

$$f'(x) = \phi(x, f(x)) \qquad (a \leq x \leq b).$$

Demostrar que semejante problema tiene a lo más una solución si hay una constante A tal que

$$|\phi(x, y_2) - \phi(x, y_1)| \leq A|y_2 - y_1|$$

siempre que $(x,y_1) \in R$, y $(x,y_2) \in R$.

Sugerencia: Aplicar el Ejercicio 26 a la diferencia de dos soluciones. Observar que este teorema de unicidad no se cumple para el problema con valores iniciales

$$y' = y^{1/2}, \qquad y(0) = 0,$$

que tiene dos soluciones: $f(x) = 0$ y $f(x) = x^2/4$. ¿Hay otras soluciones? Determinarlas.

28. Formular y demostrar un teorema de unicidad análogo, para sistemas de ecuaciones diferenciales de la forma

$$y_j' = \phi_j(x, y_1, \ldots, y_k), \qquad y_j(a) = c_j \qquad (j = 1, \ldots, k).$$

Obsérvese que puede también escribirse en la forma

$$\mathbf{y}' = \mathbf{\Phi}(x, \mathbf{y}), \qquad \mathbf{y}(a) = \mathbf{c}$$

donde $\mathbf{y} = (y_1, \ldots, y_k)$ varía sobre una celda-k, $\mathbf{\Phi}$ es el mapeo de una celda-$(k + 1)$ en el espacio-k euclidiano, cuyas componentes son las funciones $\phi_1, \ldots, \phi_k$, y $\mathbf{c}$ es el vector $(c_1, \ldots, c_k)$. Utilizar el Ejercicio 26 para funciones con valores vectoriales.

29. Particularizar el Ejercicio 28, considerando el sistema

$$y_j' = y_{j+1} \qquad (j = 1, \ldots, k - 1),$$
$$y_k' = f(x) - \sum_{j=1}^{k} g_j(x) y_j ,$$

donde f, $g_1, \ldots, g_k$ son funciones reales continuas en $[a,b]$ y deducir un teorema de unicidad para las soluciones de la ecuación

$$y^{(k)} + g_k(x)y^{(k-1)} + \cdots + g_2(x)y' + g_1(x)y = f(x),$$

con las condiciones iniciales

$$y(a) = c_1, \qquad y'(a) = c_2, \qquad \ldots, \qquad y^{(k-1)}(a) = c_k.$$

6

LA INTEGRAL DE RIEMANN-STIELTJES

El presente capítulo está basado en una definición de la integral de Riemann, que depende muy explícitamente de la estructura de orden de la recta real. En consecuencia, empezaremos estudiando la integración de las funciones reales en intervalos. En capítulos posteriores seguirán las generalizaciones a funciones de variables complejas y vectoriales en intervalos. La integración en conjuntos distintos de los intervalos, se estudia en los capítulos 10 y 11.

DEFINICIÓN Y EXISTENCIA DE LA INTEGRAL

6.1 Definición Sea $[a,b]$ un intervalo dado. Por partición P de $[a,b]$ entendemos un conjunto finito de puntos x_0, $x_1, \ldots, x_n$ donde

$$a = x_0 \le x_1 \le \cdots \le x_{n-1} \le x_n = b.$$

Escribimos

$$\Delta x_i = x_i - x_{i-1} \qquad (i = 1, \ldots, n).$$

Suponiendo, ahora, que f es una función real acotada definida en $[a,b]$. Correspondiendo a cada partición P de $[a,b]$ hacemos

$$M_i = \sup f(x) \qquad (x_{i-1} \le x \le x_i),$$
$$m_i = \inf f(x) \qquad (x_{i-1} \le x \le x_i),$$
$$U(P,f) = \sum_{i=1}^{n} M_i \, \Delta x_i,$$
$$L(P,f) = \sum_{i=1}^{n} m_i \, \Delta x_i,$$

y finalmente

(1) $$\overline{\int_a^b} f\,dx = \inf U(P,f),$$

(2) $$\underline{\int_a^b} f\,dx = \sup L(P,f),$$

donde se han considerado el sup y el inf sobre todas las particiones P de $[a,b]$. Los primeros miembros de (1) y (2) se llaman integral superior e inferior de Riemann de f sobre $[a,b]$, respectivamente.

Si las integrales superior e inferior son iguales, decimos que f es *integrable según Riemann*, sobre $[a,b]$, escribimos $f \in \mathscr{R}$ (esto es, $\mathscr{R}$ representa el conjunto de las funciones integrales según Riemann) y representamos el valor común de (1) y (2) por

(3) $$\int_a^b f\,dx,$$

o por

(4) $$\int_a^b f(x)\,dx.$$

Esta es la *integral de Riemann* de f sobre $[a,b]$. Como f es acotada, existen dos números m y M, tales que

$$m \le f(x) \le M \qquad (a \le x \le b).$$

De aquí que, para todo P,

$$m(b-a) \le L(P,f) \le U(P,f) \le M(b-a),$$

de modo que los números $L(P,f)$ y $U(P,f)$ forman un conjunto acotado. Esto demuestra que *las integrales superior e inferior están definidas* para *toda* función acotada f. La cuestión de su igualdad, y por tanto de la integrabilidad de f es más delicada. En lugar de hallarla solamente para la integral de Riemann, consideraremos inmediatamente un caso más general.

6.2 Definición Sea α una función monótona creciente en $[a,b]$ (como $\alpha(a)$ y $\alpha(b)$ son finitas, se deduce que α es acotada en $[a,b]$. Correspondiendo a cada partición P de $[a,b]$, escribimos

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$

Es claro que $\Delta\alpha_i \geq 0$. Para toda función real f acotada en $[a,b]$ escribimos

$$U(P,f,\alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i,$$

$$L(P,f,\alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i,$$

donde M_i, m_i tienen el mismo significado que en la Definición 6.1 y definimos

(5) $$\overline{\int_a^b} f\,d\alpha = \inf U(P,f,\alpha),$$

(6) $$\underline{\int_a^b} f\,d\alpha = \sup L(P,f,\alpha),$$

tomando también los inf y los sup sobre todas las particiones.

Si los primeros miembros de (5) y (6) son iguales, representamos su valor común por

(7) $$\int_a^b f\,d\alpha$$

o algunas veces por

(8) $$\int_a^b f(x)\,d\alpha(x).$$

Esta es la *integral de Riemann-Stieltjes* (o simplemente la *integral de Stieltjes*) de f respecto a α, sobre $[a,b]$.

Si existe (7), esto es si (5) y (6) son iguales, decimos que f es integral con respecto a α, en el sentido de Riemann y escribimos $f \in \mathscr{R}(\alpha)$.

Tomando $\alpha(x) = x$, se ve que la integral de Riemann es un caso particular de la integral de Riemann-Stieltjes. Merece especial mención el hecho de que en el caso general α no necesita ni ser continua.

Diremos unas palabras sobre notaciones. Preferimos (7) a (8), pues la letra x que aparece en (8) no añade nada al contenido de (7). No tiene importancia qué letra empleemos para representar la llamada «variable de integración». Por ejemplo, (8) es lo mismo que

$$\int_a^b f(y)\,d\alpha(y).$$

La integral depende de f, α, a y b, pero no de la variable de integración, que puede incluso suprimirse.

El papel que juega la variable de integración es totalmente análogo al del índice de sumación: los dos símbolos

$$\sum_{i=1}^{n} c_i, \qquad \sum_{k=1}^{n} c_k$$

expresan lo mismo, pues ambos significan $c_1 + c_2 + \cdots + c_n$.

Desde luego, no perjudica la inclusión de la variable de integración, e incluso en muchos casos es conveniente hacerlo.

Ahora, trataremos de la existencia de la integral (7). Sin mencionarlo cada vez, supondremos que f es real y acotada, y α monótona creciente en $[a,b]$; y cuando no se preste a confusión escribiremos $\int$ en lugar de $\int_a^b$.

6.3 Definición Decimos que la partición P^* es un refinamiento de P, si $P^* \supset P$ (esto es, si todo punto de P es punto de P^*). Dadas dos particiones, P_1 y P_2, decimos que P^* es su refinamiento común si $P^* = P_1 \cup P_2$.

6.4 Teorema *Si P^* es un refinamiento de P, es*

$$L(P,f,\alpha) \leq L(P^*,f,\alpha) \tag{9}$$

y

$$U(P^*,f,\alpha) \leq U(P,f,\alpha). \tag{10}$$

Demostración Para demostrar (9), supongamos primero que P^* contiene solamente un punto más que P. Sea este punto x^*, y supongamos $x_{i-1} < x^* < x_i$, donde x_{i-1} y x_i son dos puntos consecutivos de P. Hagamos

$$w_1 = \inf f(x) \qquad (x_{i-1} \leq x \leq x^*),$$
$$w_2 = \inf f(x) \qquad (x^* \leq x \leq x_i).$$

Evidentemente $w_1 \geq m_i$ y $w_2 \geq m_i$, siendo como antes

$$m_i = \inf f(x) \qquad (x_{i-1} \leq x \leq x_i).$$

Por tanto

$$\begin{aligned} &L(P^*,f,\alpha) - L(P,f,\alpha) \\ &\quad = w_1[\alpha(x^*) - \alpha(x_{i-1})] + w_2[\alpha(x_i) - \alpha(x^*)] - m_i[\alpha(x_i) - \alpha(x_{i-1})] \\ &\quad = (w_1 - m_i)[\alpha(x^*) - \alpha(x_{i-1})] + (w_2 - m_i)[\alpha(x_i) - \alpha(x^*)] \geq 0. \end{aligned}$$

Si P^* contiene k puntos más que P, repetiremos el razonamiento anterior k veces, y llegaremos a (9). La demostración de (10) es análoga.

6.5 Teorema $\underline{\int_a^b} f\,d\alpha \leq \overline{\int_a^b} f\,d\alpha.$

Demostración Sea P^* el refinamiento común de dos particiones P_1 y P_2. Por el Teorema 6.4,

$$L(P_1, f, \alpha) \leq L(P^*, f, \alpha) \leq U(P^*, f, \alpha) \leq U(P_2, f, \alpha).$$

Por tanto

(11) $$L(P_1, f, \alpha) \leq U(P_2, f, \alpha).$$

Conservando P_2 fijo, y tomando el sup sobre todo P_1, (11) da

(12) $$\underline{\int} f\,d\alpha \leq U(P_2, f, \alpha).$$

El teorema queda demostrado tomando el inf sobre todo P_2 en (12).

6.6 Teorema *$f \in \mathscr{R}(\alpha)$ en $[a,b]$ si y solo si para cada $\varepsilon > 0$ existe una partición P tal que*

(13) $$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon.$$

Demostración Para cada P tenemos

$$L(P, f, \alpha) \leq \underline{\int} f\,d\alpha \leq \overline{\int} f\,d\alpha \leq U(P, f, \alpha).$$

Así pues, (13) implica

$$0 \leq \overline{\int} f\,d\alpha - \underline{\int} f\,d\alpha < \varepsilon.$$

Por tanto, si (13) se satisface para todo $\varepsilon > 0$, tenemos

$$\overline{\int} f\,d\alpha = \underline{\int} f\,d\alpha,$$

esto es, $f \in \mathscr{R}(\alpha)$.

Inversamente, supongamos que $f \in \mathscr{R}(\alpha)$ y sea $\varepsilon > 0$ dado. Existen particiones P_1 y P_2 tales que

(14) $$U(P_2, f, \alpha) - \int f\, d\alpha < \frac{\varepsilon}{2},$$

(15) $$\int f\, d\alpha - L(P_1, f, \alpha) < \frac{\varepsilon}{2}.$$

Elijamos P como refinamiento común de P_1 y P_2. El Teorema 6.4 juntamente con (14) y (15) demuestra que

$$U(P, f, \alpha) \le U(P_2, f, \alpha) < \int f\, d\alpha + \frac{\varepsilon}{2} < L(P_1, f, \alpha) + \varepsilon \le L(P, f, \alpha) + \varepsilon,$$

de modo que se cumple (13) para esta partición P.

El Teorema 6.6 proporciona un criterio para la integrabilidad muy conveniente. Antes de aplicarlo se establecerán algunos hechos muy relacionados con este.

6.7 Teorema

(*a*) *Si se cumple* (13) *para alguna P y algún ε, entonces* (13) *se cumple también (con el mismo ε) para cada refinamiento de P.*

(*b*) *Si se cumple* (13) *para una* $P = \{x_0, \ldots, x_n\}$ *y si s_i, t_i son puntos arbitrarios en* $[x_{i-1}, x_i]$, *entonces*

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)|\, \Delta\alpha_i < \varepsilon.$$

(*c*) *Si* $f \in \mathscr{R}(\alpha)$ y la hipótesis de (*b*) *se cumple, entonces*

$$\left| \sum_{i=1}^{n} f(t_i)\, \Delta\alpha_i - \int_a^b f\, d\alpha \right| < \varepsilon.$$

Demostración El Teorema 6.4 implica (*a*). Debido a las suposiciones hechas en (*b*), se tiene que $f(s_i)$ y $f(t_i)$ están en $[m_i, M_i]$, así que $|f(s_i) - f(t_i)| \le M_i - m_i$. Entonces

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)|\, \Delta\alpha_i \le U(P, f, \alpha) - L(P, f, \alpha),$$

lo que demuestra (*b*). Las desigualdades evidentes

$$L(P, f, \alpha) \le \sum f(t_i)\, \Delta\alpha_i \le U(P, f, \alpha)$$

y

$$L(P, f, \alpha) \le \int f\, d\alpha \le U(P, f, \alpha)$$

demuestran (*c*).

6.8 Teorema *Si f es continua sobre $[a,b]$ entonces $f \in \mathscr{R}(\alpha)$ sobre $[a,b]$.*

Demostración Sea $\varepsilon > 0$ dado, y elijamos $\eta > 0$ tal que

$$[\alpha(b) - \alpha(a)]\eta < \varepsilon.$$

Como f es uniformemente continua en $[a,b]$ (Teorema 4.19), existe un $\delta >$ tal que

(16) $$|f(x) - f(t)| < \eta$$

si $|x - t| < \delta$ y $x \in [a,b]$, $t \in [a,b]$.

Si P es cualquier partición de $[a,b]$ tal que $\Delta x_i < \delta$. Entonces, (16) implica

(17) $$M_i - m_i \leq \eta \qquad (i - 1, \ldots, n)$$

Por tanto

$$U(P,f,\alpha) - L(P,f,\alpha) = \sum_{i=1}^{n} (M_i - m_i)\,\Delta\alpha_i$$
$$\leq \eta \sum_{i=1}^{n} \Delta\alpha_i = \eta[\alpha(b) - \alpha(a)] < \varepsilon.$$

Por el Teorema 6.6, $f \in \mathscr{R}(\alpha)$.

6.9 Teorema *Si f es monótona en $[a,b]$ y α es continua en $[a,b]$, $f \in \mathscr{R}(\alpha)$. (Supondremos, naturalmente, también que α es monótona.)*

Demostración Sea $\varepsilon > 0$ dado. Para cada entero positivo n, elijamos una partición P tal que

$$\Delta\alpha_i = \frac{\alpha(b) - \alpha(a)}{n} \qquad (i = 1, \ldots, n).$$

Lo que es posible, pues α es continuo (Teorema 4.23).

Supondremos que f es monótona creciente (la demostración es análoga en el otro caso). Entonces

$$M_i = f(x_i), \qquad m_i = f(x_{i-1}) \qquad (i = 1, \ldots, n),$$

de modo que

$$U(P,f,\alpha) - L(P,f,\alpha) = \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})]$$

$$= \frac{\alpha(b) - \alpha(b)}{n} \cdot [f(b) - f(a)] < \varepsilon$$

si se toma n suficientemente grande. Por el Teorema 6.6 $f \in \mathscr{R}(\alpha)$.

6.10 Teorema *Supóngase que f es acotada sobre $[a,b]$ tiene sólo un número finito de puntos de discontinuidad sobre $[a,b]$, y α es continua en cada punto para el cual f es discontinua. Entonces $f \in \mathscr{R}(\alpha)$.*

Demostración Sea $\varepsilon > 0$ dado. Hágase $M = \sup |f(x)|$, y sea E el conjunto de puntos en los cuales f es discontinua. Como E es finito y α es continua en cada punto de E, se puede cubrir E con un número finito de intervalos ajenos $[u_j, v_j] \subset [a,b]$ tales que la suma de las diferencias correspondientes $\alpha(v_j) - \alpha(u_j)$ sea menor que ε. Además, pueden reemplazarse estos intervalos de tal manera que cada punto de $E \cap (a, b)$ esté en el interior de algún $[u_j, v_j]$.

Quítense los segmentos (u_j, v_j) de $[a,b]$. El conjunto sobrante K es compacto. En consecuencia f es uniformemente continua sobre K, y existe $\delta > 0$ tal que $|f(s) - f(t)| < \varepsilon$ si $s \in K$, $t \in K$, $|s - t| < \delta$.

Ahora se forma una partición $P = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$, como sigue: cada u_j se encuentra en P. Cada v_j se encuentra en P.

Ningún punto de cualquier segmento (u_j, v_j) se encuentra en P.

Si x_{i-1} no es uno de los u_j, entonces $\Delta x_i < \delta$.

Nótese que $M_i - m_i \leq 2M$ para cada i, y que $M_i - m_i \leq \varepsilon$ a menos que x_{i-1} sea uno de los u_j. En consecuencia, de la misma forma que en la demostración del Teorema 6.8,

$$U(P, f, \alpha) - L(P, f, \alpha) \leq [\alpha(b) - \alpha(a)]\varepsilon + 2M\varepsilon.$$

Como ε es arbitrario, el Teorema 6.6 muestra que $f \in \mathscr{R}(\alpha)$.

Nota: Si f y α tienen un punto común de discontinuidad, entonces f no debe estar necesariamente en $\mathscr{R}(\alpha)$. Esto se muestra en el Ejercicio 3.

6.11 Teorema *Supongamos que $f \in \mathscr{R}(\alpha)$ en $[a,b]$, $m \leq f \leq M$, ϕ es continua en $[m,M]$, y $h(x) = \phi(f(x))$ en $[a,b]$. Entonces, $h \in \mathscr{R}(\alpha)$ en $[a,b]$.*

Demostración Elijamos $\varepsilon > 0$. Como ϕ es uniformemente continua en $[m, M]$, existe $\delta > 0$, tal que $\delta < \varepsilon$ y $|\phi(s) - \phi(t)| < \varepsilon$ si $|s - t| \leq \delta$ y $s,t \in [m, M]$.

Como $f \in \mathscr{R}(\alpha)$, hay una partición $P = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$, tal que

$$U(P, f, \alpha) - L(P, f, \alpha) < \delta^2. \tag{18}$$

Consideremos M_i y m_i con el mismo significado que en la Definición 6.1 y sean M_i^*, m_i^* los números análogos para h. Dividamos los números $1, \ldots, n$ en dos clases: $i \in A$ si $M_i - m_i < \delta$; $i \in B$ si $M_i - m_i \geq \delta$.

Para $i \in A$, la elección de δ muestra que $M_i^* - m_i^* \leq \varepsilon$.

Para $i \in B$, $M_i^* - m_i^* \leq 2K$, siendo $K = \sup |\phi(t)|$ y $m \leq t \leq M$. Por (18), tenemos

$$\delta \sum_{i \in B} \Delta\alpha_i \leq \sum_{i \in B} (M_i - m_i)\, \Delta\alpha_i < \delta^2 \tag{19}$$

de modo que $\sum_{i \in B} \Delta\alpha_i < \delta$. Se deduce que

$$\begin{aligned} U(P, h, \alpha) - L(P, h, \alpha) &= \sum_{i \in A} (M_i^* - m_i^*)\, \Delta\alpha_i + \sum_{i \in B} (M_i^* - m_i^*)\, \Delta\alpha_i \\ &\leq \varepsilon[\alpha(b) - \alpha(a)] + 2K\delta < \varepsilon[\alpha(b) - \alpha(a) + 2K]. \end{aligned}$$

Como ε era arbitrario, el Teorema 6.6 implica que $h \in \mathscr{R}(\alpha)$.

Observación: Este teorema sugiere la pregunta: ¿Qué funciones son integrables según Riemann? La contestación figura en el Teorema 11.33(b).

PROPIEDADES DE LA INTEGRAL

6.12 Teorema

(a) *Si* $f_1 \in \mathscr{R}(\alpha)$ *y* $f_2 \in \mathscr{R}(\alpha)$ *en* $[a,b]$,

$$f_1 + f_2 \in \mathscr{R}(\alpha),$$

$cf \in \mathscr{R}(\alpha)$ *para toda constante* c, *y*

$$\int_a^b (f_1 + f_2)\, d\alpha = \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha,$$

$$\int_a^b cf\, d\alpha = c \int_a^b f\, d\alpha.$$

(b) *Si* $f_1(x) \leq f_2(x)$ *en* $[a,b]$,

$$\int_a^b f_1\, d\alpha \leq \int_a^b f_2\, d\alpha.$$

(c) *Si* $f \in \mathscr{R}(\alpha)$ *en* $[a,b]$ *y* $a < c < b$, *entonces* $f \in \mathscr{R}(\alpha)$ *en* $[a,c]$ *y en* $[c,b]$ *y*

$$\int_a^c f\, d\alpha + \int_c^b f\, d\alpha = \int_a^b f\, d\alpha.$$

(d) *Si* $f \in \mathscr{R}(\alpha)$ *en* $[a,b]$ *y* $|f(x)| \leq M$ *en* $[a,b]$,

$$\left| \int_a^b f\,d\alpha \right| \leq M[\alpha(b) - \alpha(a)].$$

(*e*) *Si $f \in \mathscr{R}(\alpha_1)$ y $f \in \mathscr{R}(\alpha_2)$, entonces $f \in \mathscr{R}(\alpha_1 + \alpha_2)$ y*

$$\int_a^b f\,d(\alpha_1 + \alpha_2) = \int_a^b f\,d\alpha_1 + \int_a^b f\,d\alpha_2\,;$$

si $f \in \mathscr{R}(\alpha)$ y c es una constante positiva, será $f \in \mathscr{R}(c\alpha)$ y

$$\int_a^b f\,d(c\alpha) = c\int_a^b f\,d\alpha.$$

Demostración Si $f = f_1 + f_2$ y P es alguna partición de $[a,b]$, tenemos que

(20) $$L(P,f_1,\alpha) + L(P,f_2,\alpha) \leq L(P,f,\alpha) \leq U(P,f,\alpha) \leq U(P,f_1,\alpha) + U(P,f_2,\alpha).$$

Si $f_1 \in \mathscr{R}(\alpha)$ y $f_2 \in \mathscr{R}(\alpha)$, sea $\varepsilon > 0$ dado. Existen particiones P_j ($j = 1, 2$), tales que

$$U(P_j,f_j,\alpha) - L(P_j,f_j,\alpha) < \varepsilon.$$

Esas desigualdades subsisten si se sustituyen P_1 y P_2 por su refinamiento común P. Entonces (20) implica

$$U(P,f,\alpha) - L(P,f,\alpha) < 2\varepsilon,$$

lo que demuestra que $f \in \mathscr{R}(\alpha)$.

Con este mismo P, tenemos

$$U(P,f_j,\alpha) < \int f_j\,d\alpha + \varepsilon \qquad (j = 1, 2);$$

por lo que (20) implica

$$\int f\,d\alpha \leq U(P,f,\alpha) < \int f_1\,d\alpha + \int f_2\,d\alpha + 2\varepsilon.$$

Como ε era arbitrario, deducimos que

(21) $$\int f\,d\alpha \leq \int f_1\,d\alpha + \int f_2\,d\alpha.$$

Si sustituimos en (21) f_1 y f_2 por $-f_1$ y $-f_2$, se invierte la desigualdad, y queda demostrada la igualdad.

Las demostraciones de las otras afirmaciones del Teorema 6.12 son tan parecidas que omitimos los detalles. En el apartado (c) podemos limitarnos pasando a refinamientos a las particiones que contienen al punto c, al aproximar $\int f\,d\alpha$.

6.13 Teorema *Si $f \in \mathscr{R}(\alpha)$ y $g \in \mathscr{R}(\alpha)$ en $[a,b]$,*

(a) $fg \in \mathscr{R}(\alpha)$;

(b) $|f| \in \mathscr{R}(\alpha)$ y $\left|\int_a^b f\,d\alpha\right| \le \int_a^b |f|\,d\alpha$.

Demostración Si tomamos $\phi(t) = t^2$, el Teorema 6.11 demuestra que $f^2 \in \mathscr{R}(\alpha)$ si $f \in \mathscr{R}(\alpha)$. La identidad

$$4fg = (f+g)^2 - (f-g)^2$$

completa la demostración de (a).

Si tomamos $\phi(t) = |t|$, el Teorema 6.11 demuestra, de igual modo, que $|f| \in \mathscr{R}(\alpha)$. Elijamos $c = \pm 1$, de forma que

$$c\int f\,d\alpha \ge 0.$$

Será

$$\left|\int f\,d\alpha\right| = c\int f\,d\alpha = \int cf\,d\alpha \le \int |f|\,d\alpha,$$

pues $cf \le |f|$.

6.14 Definición La *función escalón unitario* I se define como

$$I(x) = \begin{cases} 0 & (x \le 0), \\ 1 & (x > 0). \end{cases}$$

6.15 Teorema *Si $a < s < b$ es acotada sobre $[a,b]$, f es continua en s, y $\alpha(x) = I(x-s)$, entonces*

$$\int_a^b f\,d\alpha = f(s).$$

Demostración Considérense particiones $P = \{x_0, x_1, x_2, x_3\}$, donde $x_0 = a$, y $x_1 = s < x_2 < x_3 = b$. Entonces

$$U(P,f,\alpha) = M_2, \qquad L(P,f,\alpha) = m_2.$$

Debido a que f es continua en s, se ve que M_2 y m_2 converge hacia $f(s)$ cuando $x_2 \to s$.

6.16 Teorema *Supóngase que $c_n \geq 0$ para 1, 2, 3,..., Σc_n converge, $\{s_n\}$ es una sucesión de puntos distintos en (a,b), y*

(22) $$\alpha(x) = \sum_{n=1}^{\infty} c_n I(x - s_n).$$

Sea f continua sobre $[a,b]$. Entonces

(23) $$\int_a^b f\,d\alpha = \sum_{n=1}^{\infty} c_n f(s_n).$$

Demostración La prueba de comparación muestra que la serie (22) converge para cada x. Su suma $\alpha(x)$ es evidentemente monótona y $\alpha(a) = 0$, $\alpha(b) = \Sigma c_n$. (Este es el tipo de función que se presentó en la Observación 4.31.)

Sea $\varepsilon > 0$ conocido, y elíjase N de manera que

$$\sum_{N+1}^{\infty} c_n < \varepsilon.$$

Haciendo

$$\alpha_1(x) = \sum_{n=1}^{N} c_n I(x - s_n), \qquad \alpha_2(x) = \sum_{N+1}^{\infty} c_n I(x - s_n).$$

De los Teoremas 6.12 y 6.15 se tiene

(24) $$\int_a^b f\,d\alpha_1 = \sum_{i=1}^{N} c_n f(s_n).$$

Como $\alpha_2(b) - \alpha_2(a) < \varepsilon$, entonces

(25) $$\left| \int_a^b f\,d\alpha_2 \right| \leq M\varepsilon,$$

donde $M = \sup|f(x)|$. Debido a que $\alpha = \alpha_1 + \alpha_2$, se deduce de (24) y (25) que

(26) $$\left| \int_a^b f\,d\alpha - \sum_{i=1}^{N} c_n f(s_n) \right| \leq M\varepsilon.$$

Si se hace $N \to \infty$, se obtiene (23).

6.17 Teorema *Si se supone que α crece monótonamente y $\alpha' \in \mathscr{R}$ sobre $[a,b]$. Sea f una función real acotada sobre $[a,b]$.*

Entonces $f \in \mathscr{R}(\alpha)$ si, y solo si $f\alpha' \in \mathscr{R}$. En este caso

(27) $$\int_a^b f\,d\alpha = \int_a^b f(x)\alpha'(x)\,dx.$$

Demostración Sea $\varepsilon > 0$ dado y aplíquese el Teorema 6.6 a α': Existe una partición $P = \{x_0, \ldots, x_n\}$ de $[a, b]$ tal que

(28) $$U(P, \alpha') - L(P, \alpha') < \varepsilon.$$

El teorema del valor medio suministra puntos $t_i \in [x_{i-1}, x_i]$ tales que

$$\Delta\alpha_i = \alpha'(t_i)\,\Delta x_i$$

para $i = 1, \ldots, n$. Si $s_i \in [x_{i-1}, x_i]$, entonces

(29) $$\sum_{i=1}^{n} |\alpha'(s_i) - \alpha'(t_i)|\,\Delta x_i < \varepsilon,$$

de (28) y el Teorema 6.7(b). Hacer ahora $M = \sup|f(x)|$. Ya que

$$\sum_{i=1}^{n} f(s_i)\,\Delta\alpha_i = \sum_{i=1}^{n} f(s_i)\alpha'(t_i)\,\Delta x_i$$

y se deduce de (29) que

(30) $$\left|\sum_{i=1}^{n} f(s_i)\,\Delta\alpha_i - \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\,\Delta x_i\right| \leq M\varepsilon.$$

En particular,

$$\sum_{i=1}^{n} f(s_i)\,\Delta\alpha_i \leq U(P, f\alpha') + M\varepsilon,$$

para todas las elecciones de $s_i \in [x_{i-1}, x_i]$, de manera que

$$U(P, f, \alpha) \leq U(P, f\alpha') + M\varepsilon.$$

De (30), y con el mismo argumento se obtiene

$$U(P, f\alpha') \leq U(P, f, \alpha) + M\varepsilon.$$

Entonces

(31) $$|U(P, f, \alpha) - U(P, f\alpha')| \leq M\varepsilon.$$

Ahora nótese que, (28) sigue siendo cierto si P se reemplaza por

cualquier refinamiento. De aquí que (31) sigue siendo también cierto. Se concluye que

$$\left| \overline{\int_a^b} f\,d\alpha - \overline{\int_a^b} f(x)\alpha'(x)\,dx \right| \le M\varepsilon.$$

Pero ε es arbitrario. En consecuencia,

(32) $$\overline{\int_a^b} f\,d\alpha = \overline{\int_a^b} f(x)\alpha'(x)\,dx,$$

para *cualquier* f acotada. La igualdad de las integrales inferiores se deduce de (30) exactamente de la misma forma. De aquí se deduce el teorema.

6.18 Observación Los dos teoremas anteriores ilustran la generalidad y flexibilidad que son inherentes en el proceso de integración de Stieltjes. Si α es una función escalón pura [este es el nombre que con frecuencia se da a las funciones de la forma (22)], la integral se reduce a una serie finita o infinita. Si α tiene derivada integrable, la integral se reduce a una integral de Riemann ordinaria. Esto hace posible en la mayoría de los casos estudiar series e integrales en forma simultánea, en vez de separadamente.

Considérese un ejemplo físico para ilustrar lo anterior. El momento de inercia de un alambre recto de longitud unitaria a través de uno de sus extremos y con respecto a un eje, que forma un ángulo recto con el alambre, es

(33) $$\int_0^1 x^2\,dm$$

en donde $m(x)$ es la masa que se tiene en el intervalo $[0, x]$. Si se considera que el alambre tiene densidad continua ρ, esto es, si $m'(x) = \rho(x)$, entonces (33) se vuelve

(34) $$\int_0^1 x^2\,\rho(x)\,dx.$$

Por otro lado, si el alambre está compuesto de masas m_i concentradas en puntos x_i, (33) se convierte en

(35) $$\sum_i x_i^2\,m_i.$$

Es por esto que (33) contiene como casos especiales a (34) y (35), pero también contiene mucho más; por ejemplo, el caso en el cual m es continua, pero no diferenciable en todas partes.

6.19 Teorema (cambio de variable) *Supóngase que φ es una función continua estrictamente creciente que mapea un intervalo $[A,B]$ sobre $[a,b]$. Su-*

póngase también que α es monótona creciente sobre $[a,b]$ y $f \in \mathscr{R}(\alpha)$ sobre $[a,b]$. Si se define β y g sobre $[A,B]$ por medio de

(36) $$\beta(y) = \alpha(\varphi(y)), \qquad g(y) = f(\varphi(y)).$$

Entonces $g \in \mathscr{R}(\beta)$ y

(37) $$\int_A^B g\, d\beta = \int_a^b f\, d\alpha.$$

Demostración A cada partición $P = \{x_0, \ldots, x_n\}$ de $[a,b]$ le corresponde una partición $Q = \{y_0, \ldots, y_n\}$ de $[A,B]$, de tal manera que $x_i = \varphi(y_i)$. Todas las particiones de $[A,B]$ se obtienen de esta forma. Como los valores tomados por f sobre $[x_{i-1}, x_i]$ son exactamente los mismos que los tomados por g sobre $[y_{i-1}, y_i]$, se ve que

(38) $$U(Q, g, \beta) = U(P, f, \alpha), \qquad L(Q, g, \beta) = L(P, f, \alpha).$$

Debido a que $f \in \mathscr{R}(\alpha)$, P puede elegirse de tal manera que $U(P,f,\alpha)$ y $L(P,f,\alpha)$ estén próximas a $\int f\, d\alpha$. De aquí (38) combinada con el Teorema 6.6 muestra que $g \in \mathscr{R}(\beta)$ y que (37) se cumple. Esto completa la demostración.

Nótese el siguiente caso especial:

Tomando $\alpha(x) = x$. Entonces $\beta = \varphi$. Suponiendo $\varphi' \in \mathscr{R}$ sobre $[A,B]$. Si se aplica el Teorema 6.17 al miembro izquierdo de (37), se obtiene

(39) $$\int_a^b f(x)\, dx = \int_A^B f(\varphi(y))\varphi'(y)\, dy.$$

INTEGRACIÓN Y DIFERENCIACIÓN

En esta sección continuamos limitándonos a las funciones reales. Demostraremos que la integración y la diferenciación son, en cierto sentido, operaciones inversas.

6.20 Teorema *Sea $f \in \mathscr{R}$ en $[a,b]$. Para $a \le x \le b$, hagamos*

$$F(x) = \int_a^x f(t)\, dt.$$

En estas condiciones, F es continua en $[a,b]$; además, si f es continua en un punto x_0 de $[a,b]$, F es diferenciable en x_0, y

$$F'(x_0) = f(x_0).$$

Demostración Como $f \in \mathscr{R}$, es acotada. Supongamos $|f(t)| \leq M$ para $a \leq t \leq b$. Si $a \leq x < y \leq b$, será

$$|F(y) - F(x)| = \left|\int_x^y f(t)\, dt\right| \leq M(y - x),$$

por el Teorema 6.12(*c*) y (*d*). Dado $\varepsilon > 0$, vemos que

$$|F(y) - F(x)| < \varepsilon,$$

con tal que $|y - x| < \varepsilon/M$. Esto demuestra la continuidad (en realidad la continuidad uniforme) de F.

Supongamos, ahora, que f es continua en x_0. Dado $\varepsilon > 0$, elijamos $\delta > 0$, tal que

$$|f(t) - f(x_0)| < \varepsilon$$

si $|t - x_0| < \delta$ y $a \leq t \leq b$. Por tanto, si

$$x_0 - \delta < s \leq x_0 \leq t < x_0 + \delta \qquad \text{y} \qquad a \leq s < t \leq b,$$

tenemos, por el Teorema 6.12(*d*)

$$\left|\frac{F(t) - F(s)}{t - s} - f(x_0)\right| = \left|\frac{1}{t - s}\int_s^t [f(u) - f(x_0)]\, du\right| < \varepsilon.$$

Se deduce así, que $F'(x_0) = f(x_0)$

6.21 El teorema fundamental del cálculo. *Si $f \in \mathscr{R}$ sobre $[a,b]$ y si existe una función diferenciable F sobre $[a,b]$ tal que $F' = f$, entonces*

$$\int_a^b f(x)\, dx = F(b) - F(a).$$

Demostración Dado $\varepsilon > 0$, elíjase una partición $P = \{x_0, \ldots, x_n\}$ de $[a,b]$, de tal manera que $U(P,f) - L(P,f) < \varepsilon$. El teorema del valor medio proporciona los puntos $t_i \in [x_{i-1}, x_i]$ de tal manera que

$$F(x_i) - F(x_{i-1}) = f(t_i)\, \Delta x_i$$

para $i = 1, \ldots, n$. Entonces

$$\sum_{i=1}^n f(t_i)\, \Delta x_i = F(b) - F(a).$$

Y del Teorema 6.7(c) se deduce ahora que

$$\left|F(b) - F(a) - \int_a^b f(x)\, dx\right| < \varepsilon.$$

Por que esto se verifica para cada $\varepsilon > 0$, la demostración queda concluida.

6.22 Teorema (integración por partes) *Si F y G son funciones diferenciables sobre $[a,b]$, $F' = f \in \mathscr{R}$, y $G' = g \in \mathscr{R}$. Entonces*

$$\int_a^b F(x)g(x)\, dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)\, dx.$$

Demostración Haciendo $H(x) = F(x)G(x)$ y aplicando el Teorema 6.21 a H y su derivada. Se ve claramente que $H' \in \mathscr{R}$, debido al Teorema 6.13.

INTEGRACIÓN DE FUNCIONES VECTORIALES

6.23 Definición Sean $f_1, \ldots, f_k$ funciones reales en $[a,b]$ y $\mathbf{f} = (f_1, \ldots, f_k)$ la correspondiente aplicación de $[a,b]$ en R^k. Si α es monótona creciente en $[a,b]$, decir que $\mathbf{f} \in \mathscr{R}(\alpha)$ significa que $f_j \in \mathscr{R}(\alpha)$ para $j = 1, \ldots, k$. En este caso, definimos

$$\int_a^b \mathbf{f}\, d\alpha = \left(\int_a^b f_1\, d\alpha, \ldots, \int_a^b f_k\, d\alpha\right).$$

En otras palabras, $\int \mathbf{f}\, d\alpha$ es el punto en R^k cuya coordenada j-ésima es $\int f_j\, d\alpha$.

Es claro que los apartados (a), (c) y (e) del Teorema 6.12 son válidos para estas integrales con valores vectoriales; no hacemos más que aplicar los resultados primitivos a cada coordenada. Lo mismo es cierto respecto al Teorema 6.17, 6.20 y 6.21. Como aclaración, enunciamos el análogo al Teorema 6.21.

6.24 Teorema *Si $\mathbf{f}$ y $\mathbf{F}$ aplican $[a,b]$ en R^k, si $\mathbf{f} \in \mathscr{R}$ en $[a,b]$ y $\mathbf{F}' = \mathbf{f}$, entonces*

$$\int_a^b \mathbf{f}(t)\, dt = \mathbf{F}(b) - \mathbf{F}(a).$$

El análogo al Teorema 6.13(b) presenta, sin embargo, algún aspecto nuevo, al menos en la demostración:

6.25 Teorema *Si $\mathbf{f}$ aplica $[a,b]$ en R^k y $\mathbf{f} \in \mathscr{R}(\alpha)$ para alguna función monótona creciente α en $[a,b]$, entonces $|\mathbf{f}| \in \mathscr{R}(\alpha)$ y*

(40) $$\left|\int_a^b \mathbf{f}\, d\alpha\right| \le \int_a^b |\mathbf{f}|\, d\alpha.$$

Demostración Si $f_1, \ldots, f_k$ son las componentes de $\mathbf{f}$,

$$|\mathbf{f}| = (f_1^2 + \cdots + f_k^2)^{1/2}. \tag{41}$$

Por el Teorema 6.11, cada una de las funciones f_i^2 pertenece a $\mathscr{R}(\alpha)$, por lo que también su suma. Como x^2 es una función continua de x, el Teorema 4.17 demuestra que la función raíz cuadrada es continua en $[0, M]$, para todo número real M. Si aplicamos una vez más el Teorema 6.11, (41) demuestra que $|\mathbf{f}| \in \mathscr{R}(\alpha)$.

Para probar (40), hagamos $\mathbf{y} = (y_1, \ldots, y_k)$ donde $y_j = \int f_j \, d\alpha$. Será $\mathbf{y} = \int \mathbf{f} \, d\alpha$, y

$$|\mathbf{y}|^2 = \sum y_i^2 = \sum y_j \int f_j \, d\alpha = \int \left(\sum y_j f_j\right) d\alpha.$$

Por la desigualdad de Schwarz,

$$\sum y_j f_j(t) \leq |\mathbf{y}| \, |\mathbf{f}(t)| \qquad (a \leq t \leq b); \tag{42}$$

por lo que el Teorema 6.12(b) implica

$$|\mathbf{y}|^2 \leq |\mathbf{y}| \int |\mathbf{f}| \, d\alpha. \tag{43}$$

Si $\mathbf{y} = \mathbf{0}$, (40) es elemental. Si $\mathbf{y} \neq \mathbf{0}$, la división de (43) por $|\mathbf{y}|$ da (40).

CURVAS RECTIFICABLES

Terminamos este capítulo con un tema, de interés en geometría, que proporciona una aplicación de algo de la teoría precedente. El caso $k = 2$ (esto es, el caso de las curvas planas) es de importancia considerable en el estudio de las funciones analíticas de variable compleja.

6.26 Definición Una aplicación continua γ de un intervalo $[a,b]$ en R^k se llama *curva* en R^k. Para hacer notar el intervalo del parámetro $[a,b]$, se dice también que γ es una curva sobre $[a,b]$.

Si γ es uno-a-uno, γ se llama *arco*.

Si $\gamma(a) = \gamma(b)$, se dice que es una *curva cerrada*.

Debe observarse que se ha definido una curva como una *aplicación*, no como un conjunto de puntos. Por supuesto que cada curva γ en R^k tiene asociado un subconjunto de R^k, es decir el rango de γ, pero que diferentes curvas pueden tener el mismo rango.

A cada partición $P = \{x_0, \ldots, x_n\}$ de $[a,b]$ y a toda curva γ sobre $[a,b]$ se le asocia el número

$$\Lambda(P, \gamma) = \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})|.$$

El término i-ésimo en esta suma es la distancia (en R^k) entre los puntos $\gamma(x_{i-1})$ y $\gamma(x_i)$. En consecuencia $\Lambda(P,\gamma)$ es la longitud de una trayectoria poligonal con vértices en $\gamma(x_0)$, $\gamma(x_1), \ldots, \gamma(x_n)$, conservando este orden. Conforme la partición se hace más fina, este polígono se aproxima al rango de γ cada vez más. Esto hace razonable definir la *longitud de* γ como

$$\Lambda(\gamma) = \sup \Lambda(P, \gamma),$$

donde el supremum se toma sobre todas las particiones de d$[a,b]$.

Si $\Lambda(\gamma) < \infty$, se dice que γ es *rectificable.*

En algunos casos, $\Lambda(\gamma)$ se da como una integral de Riemann. Se demostrará esto para curvas *continuamente diferenciables*, es decir, para curvas γ cuya derivada γ' es continua.

6.27 Teorema *Si* γ' *es continua sobre* $[a,b]$, *entonces* γ *es rectificable, y*

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)|\, dt.$$

Demostración Si $a \leq x_{i-1} < x_i \leq b$, entonces

$$|\gamma(x_i) - \gamma(x_{i-1})| = \left|\int_{x_{i-1}}^{x_i} \gamma'(t)\, dt\right| \leq \int_{x_{i-1}}^{x_i} |\gamma'(t)|\, dt.$$

Por consiguiente

$$\Lambda(P, \gamma) \leq \int_a^b |\gamma'(t)|\, dt$$

para cada partición P de $[a,b]$. En consecuencia,

$$\Lambda(\gamma) \leq \int_a^b |\gamma'(t)|\, dt.$$

Para demostrar la desigualdad opuesta, sea $\varepsilon > 0$ conocido. Como γ' es uniformemente continua sobre $[a,b]$, existe $\delta > 0$ tal que

$$|\gamma'(s) - \gamma'(t)| < \varepsilon \qquad \text{if } |s - t| < \delta.$$

Sea $P = \{x_0, \ldots, x_n\}$ una partición de $[a,b]$, con $\Delta x_i < \delta$ para toda i. Si $x_{i-1} \leq t \leq x_i$, se deduce que

$$|\gamma'(t)| \leq |\gamma'(x_i)| + \varepsilon.$$

Por tanto

$$\begin{aligned}\int_{x_{i-1}}^{x_i} |\gamma'(t)|\, dt &\le |\gamma'(x_i)|\,\Delta x_i + \varepsilon\,\Delta x_i \\ &= \left|\int_{x_{i-1}}^{x_i} [\gamma'(t) + \gamma'(x_i) - \gamma'(t)]\, dt\right| + \varepsilon\,\Delta x_i \\ &\le \left|\int_{x_{i-1}}^{x_i} \gamma'(t)\, dt\right| + \left|\int_{x_{i-1}}^{x_i} [\gamma'(x_i) - \gamma'(t)]\, dt\right| + \varepsilon\,\Delta x_i \\ &\le |\gamma(x_i) - \gamma(x_{i-1})| + 2\varepsilon\,\Delta x_i .\end{aligned}$$

Si se adicionan estas desigualdades, se obtiene

$$\begin{aligned}\int_a^b |\gamma'(t)|\, dt &\le \Lambda(P, \gamma) + 2\varepsilon(b-a) \\ &\le \Lambda(\gamma) + 2\varepsilon(b-a).\end{aligned}$$

Debido a que ε era arbitrario,

$$\int_a^b |\gamma'(t)|\, dt \le \Lambda(\gamma).$$

Esto completa la demostración.

EJERCICIOS

1. Suponer que α es creciente en $[a,b]$; $a \le x_0 \le b$; α es continua en x_0; $f(x_0) = 1$, y $f(x) = 0$ si $x \ne x_0$. Demostrar que $f \in \mathscr{R}(\alpha)$ y que $\int f\, d\alpha = 0$.
2. Suponer que $f \ge 0$; f es continua en $[a,b]$, y $\int_a^b f(x)\, dx = 0$. Demostrar que $f(x) = 0$ para todo $x \in [a,b]$. (Comparar con el Ejercicio 1.)
3. Definir tres funciones $\beta_1, \beta_2, \beta_3$ como sigue: $\beta_j(x) = 0$ si $x < 0$; $\beta_j(x) = 1$ si $x > 0$ para $j = 1, 2, 3$; y $\beta_1(0) = 0$; $\beta_2(0) = 1$; $\beta_3(0) = \frac{1}{2}$. Sea f una función acotada en $[-1, 1]$.

   (*a*) Demostrar que $f \in \mathscr{R}(\beta_1)$ si, y solo si $f(0+) = f(0)$, y que

   $$\int f\, d\beta_1 = f(0). \text{ en este caso.}$$

   (*b*) Plantear y demostrar un resultado similar para β_2.

   (*c*) Demostrar que $f \in \mathscr{R}(\beta_3)$ si, y solo si f es continua en 0.

   (*d*) Si f es continua en 0, demostrar que

   $$\int f\, d\beta_1 = \int f\, d\beta_2 = \int f\, d\beta_3 = f(0).$$
4. Si $f(x) = 0$ para todo número irracional x, y $f(x) = 1$ para todo racional x, demostrar que $f \notin \mathscr{R}$ en $[a,b]$ para un $a < b$.

5. Supóngase que f es una función real acotada sobre $[a,b]$, y que $f^2 \in \mathscr{R}$ sobre $[a,b]$. ¿Se puede deducir que $f \in \mathscr{R}$? Si se supone que $f^3 \in \mathscr{R}$, ¿cambia la respuesta?

6. Sea el conjunto de Cantor P construido en la sección 2.44. Sea f una función real acotada sobre $[0, 1]$ continua en cada punto que esté fuera de P. Demostrar que $f \in \mathscr{R}$ sobre $[0, 1]$. *Sugerencia*: P puede cubrirse con un número finito de segmentos cuya longitud total pueda hacerse tan pequeña como se desee. Proceder de la misma forma que en el Teorema 6.10.

7. Supóngase que f es una función real definida sobre $(0, 1]$ y que $f \in \mathscr{R}$ sobre $[c, 1]$ para cada $c > 0$. Se define

$$\int_0^1 f(x)\, dx = \lim_{c \to 0} \int_c^1 f(x)\, dx$$

si el límite existe (y es finito).

(*a*) Si $f \in \mathscr{R}$ sobre $[0, 1]$, mostrar que esta definición de la integral coincide con la definición antigua.

(*b*) Construir una función f tal que el límite anterior exista, aunque no exista cuando $|f|$ se reemplace por f.

8. Supóngase que $f \in \mathscr{R}$ sobre $[a,b]$ para cada $b > a$ con a fija. Se define

$$\int_a^\infty f(x)\, dx = \lim_{b \to \infty} \int_a^b f(x)\, dx$$

si el límite existe (y es finito). En este caso se dice que la integral de la izquierda *converge*. Si ésta también converge después de haber reemplazado f por $|f|$, entonces se dice que converge *absolutamente*.

Supóngase ahora que $f(x) \geq 0$ y que f es monótona decreciente sobre $[1, \infty)$. Demostrar que

$$\int_1^\infty f(x)\, dx$$

converge si, y solo si

$$\sum_{n=1}^\infty f(n)$$

converge. (Este es el llamado "criterio de la integral" para la convergencia de series.)

9. Mostrar que algunas veces puede aplicarse la integración por partes a las integrales "impropias" que se definieron en los Ejercicios 7 y 8. (Formular un teorema estableciendo las hipótesis apropiadas y demostrarlo.) Por ejemplo mostrar que

$$\int_0^\infty \frac{\cos x}{1+x}\, dx = \int_0^\infty \frac{\operatorname{sen} x}{(1+x)^2}\, dx.$$

Mostrar que una de estas integrales converge *absolutamente*, pero la otra no.

10. Sean p y q dos números reales positivos tales que

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Demostrar lo siguiente:
(*a*) Si $u \geq 0$ y $v \geq 0$, entonces

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q}.$$

La igualdad es cierta si, y solo si $u^p = v^q$.
(*b*) Si $f \in \mathscr{R}(\alpha)$, $g \in \mathscr{R}(\alpha)$, $f \geq 0$, $g \geq 0$, y

$$\int_a^b f^p \, d\alpha = 1 = \int_a^b g^q \, d\alpha,$$

entonces

$$\int_a^b fg \, d\alpha \leq 1.$$

(*c*) Si f y g son funciones complejas en $\mathscr{R}(\alpha)$, entonces

$$\left| \int_a^b fg \, d\alpha \right| \leq \left\{ \int_a^b |f|^p \, d\alpha \right\}^{1/p} \left\{ \int_a^b |g|^q \, d\alpha \right\}^{1/q}.$$

Esta es la *desigualdad de Hölder*. Cuando $p = q = 2$, se llama comúnmente desigualdad de Schwarz. (Nótese que el Teorema 1.35 es un caso muy especial de ésta.)
(*d*) Mostrar que la desigualdad de Hölder también es verdadera para las integrales "impropias" que se describieron en los Ejercicios 7 y 8.

11. Sea α una función creciente fija sobre $[a,b]$. Si $u \in \mathscr{R}(\alpha)$ se define

$$\|u\|_2 = \left\{ \int_a^b |u|^2 \, d\alpha \right\}^{1/2}.$$

Supóngase que f, g, $h \in \mathscr{R}(\alpha)$, y demuéstrese la desigualdad del triángulo

$$\|f - h\|_2 \leq \|f - g\|_2 + \|g - h\|_2$$

como una consecuencia de la desigualdad de Schwarz, de la misma manera que en la demostración del Teorema 1.37.

12. Con las notaciones del Ejercicio 11, supóngase que $f \in \mathscr{R}(\alpha)$ y $\varepsilon > 0$. Demostrar que existe una función continua g sobre $[a,b]$ tal que $\|f - g\|_2 < \varepsilon$.

Sugerencia: Sea $P = \{x_0, \ldots, x_n\}$ una partición adecuada de $[a,b]$, y defínase

$$g(t) = \frac{x_i - t}{\Delta x_i} f(x_{i-1}) + \frac{t - x_{i-1}}{\Delta x_i} f(x_i)$$

si $x_{i-1} \leq t \leq x_i$.

13. Si se define

$$f(x) = \int_x^{x+1} \operatorname{sen}(t^2)\, dt.$$

(*a*) Demostrar que $|f(x)| < 1/x$ si $x > 0$.

Sugerencia: Hacer $t^2 = u$ e integrar por partes, para mostrar que $f(x)$ es igual a

$$\frac{\cos(x^2)}{2x} - \frac{\cos[(x+1)^2]}{2(x+1)} - \int_{x^2}^{(x+1)^2} \frac{\cos u}{4u^{3/2}}\, du.$$

Reemplazar $\cos u$ por -1.

(*b*) Demostrar que

$$2xf(x) = \cos(x^2) - \cos[(x+1)^2] + r(x)$$

en donde $|r(x)| < c/x$ y c es una constante.

(*c*) Encontrar los límites superior e inferior de $xf(x)$, cuando $x \to \infty$.

(*d*) ¿Converge $\int_0^\infty \operatorname{sen}(t^2)\, dt$?

14. Si

$$f(x) = \int_x^{x+1} \operatorname{sen}(e^t)\, dt.$$

Mostrar que

$$e^x |f(x)| < 2$$

y que

$$e^x f(x) = \cos(e^x) - e^{-1}\cos(e^{x+1}) + r(x),$$

en donde $|r(x)| < Ce^{-x}$, para alguna constante C.

15. Supóngase que f es real, y continuamente diferenciable sobre $[a,b]$, $f(a) = f(b) = 0$, y

$$\int_a^b f^2(x)\, dx = 1.$$

Demostrar que

$$\int_a^b xf(x)f'(x)\, dx = -\tfrac{1}{2}$$

y también que

$$\int_a^b [f'(x)]^2\, dx \cdot \int_a^b x^2 f^2(x)\, dx > \tfrac{1}{4}.$$

16. Para $1 < s < \infty$, se define

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

(Ésta es la función zeta de Riemann, que es muy importante en el estudio de la distribución de los números primos.) Demostrar que

$$(a)\ \zeta(s) = s\int_1^\infty \frac{[x]}{x^{s+1}}\,dx$$

y que

$$(b)\ \zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{x-[x]}{x^{s+1}}\,dx,$$

en donde $[x]$ representa el entero mayor $\leq x$.

Demostrar que la integral del apartado (b) converge para todo $s > 0$.

Sugerencia: Para demostrar (a), calcular la diferencia entre la integral sobre $[1, N]$ y la n-ésima suma parcial de la serie que define a $\zeta(s)$.

17. Supóngase que α crece monótonamente sobre $[a,b]$, que g es continua, y $g(x) = G'(x)$ para $a \leq x \leq b$. Demostrar que

$$\int_a^b \alpha(x)g(x)\,dx = G(b)\alpha(b) - G(a)\alpha(a) - \int_a^b G\,d\alpha.$$

Sugerencia: Tomar g real, sin perder generalidad. Dado $P = \{x_0, x_1, \ldots, x_n\}$, elegir $t_i \in (x_{i-1}, x_i)$ de tal manera que $g(t_i)\,\Delta x_i = G(x_i) - G(x_{i-1})$. Mostrar que

$$\sum_{i=1}^n \alpha(x_i)g(t_i)\,\Delta x_i = G(b)\alpha(b) - G(a)\alpha(a) - \sum_{i=1}^n G(x_{i-1})\,\Delta\alpha_i.$$

18. Sean γ_1, γ_2, γ_3, curvas en el plano complejo, definidas en $[0, 2\pi]$ por

$$\gamma_1(t) = e^{it}, \qquad \gamma_2(t) = e^{2it}, \qquad \gamma_3(t) = e^{2\pi it\,\mathrm{sen}\,(1/t)}.$$

Demostrar que estas tres curvas tienen el mismo rango que γ_1 y γ_2 son rectificables, que la longitud de γ_1 es 2π, que la de γ_2 es 4π y que γ_3 no es rectificable.

19. Sea γ_1 una curva en R^k, definida en $[a,b]$; sea ϕ un mapeo 1-1 continuo de $[c,d]$ sobre $[a,b]$, tal que $\phi(c) = a$, y definamos $\gamma_2(s) = \gamma_1(\phi(s))$. Demostrar que γ_2 es un arco, una curva cerrada simple o una curva rectificable, si y solo si es cierto lo mismo para γ_1. Demostrar que γ_2 y γ_1 tienen la misma longitud.

7

SUCESIONES Y SERIES DE FUNCIONES

En el presente capítulo limitaremos nuestra atención a las funciones de variables complejas (en ellas incluiremos, naturalmente, las de valores reales) aunque muchos de los teoremas y demostraciones que siguen se amplian sin dificultad a las funciones vectoriales, e incluso a los mapeos en espacios métricos en general. Elegimos el trabajar en este marco reducido para fijar la atención en los aspectos más importantes de los problemas que se presentan cuando se varía el orden en los procesos de límites.

DISCUSIÓN DEL PROBLEMA PRINCIPAL

7.1 Definición Supongamos que $\{f_n\}$, con $n = 1, 2, 3, \ldots$, es una sucesión de funciones definidas en un conjunto E, y que la sucesión de números $\{f_n(x)\}$ converge para todo $x \in E$. Podemos definir una función f por

(1) $$f(x) = \lim_{n \to \infty} f_n(x) \qquad (x \in E).$$

En estas circunstancias, decimos que $\{f_n\}$ converge en E y que f es el *límite* o la *función límite, de* $\{f_n\}$. A veces, utilizaremos una terminología más expresiva y diremos que «$\{f_n\}$ converge hacia f *puntualmente* en E», o bien, «en cada punto de E», si se cumple (1). Del mismo modo, si $\Sigma f_n(x)$ converge para todo $x \in E$ y definimos

(2) $$f(x) = \sum_{n=1}^{\infty} f_n(x) \qquad (x \in E),$$

a la función f se le llama *suma* de la serie Σf_n.

El principal problema que se presenta es el de determinar qué propiedades de las funciones se conservan con las operaciones de límites (1) y (2). Por ejemplo, si las funciones f_n son continuas, o diferenciables o integrables, ¿sucede lo mismo con la función límite? ¿Cuáles son las relaciones entre f_n' y f' o entre las integrales de f_n y la de f?

Decir que f es continua en x significa que

$$\lim_{t \to x} f(t) = f(x).$$

Por tanto, preguntar si el límite de una sucesión de funciones continuas es continuo, es lo mismo que preguntar si

(3) $$\lim_{t \to x} \lim_{n \to \infty} f_n(t) = \lim_{n \to \infty} \lim_{t \to x} f_n(t),$$

esto es, si no importa el orden en que se aplica el proceso de límites. En el primer miembro de (3), hacemos primero $n \to \infty$, y luego, $t \to x$; en el segundo, primero $t \to x$, y luego, $n \to \infty$.

Vamos a demostrar, por medio de ejemplos que, en general, no se puede variar el orden en los procesos de límites sin afectar al resultado. Después demostraremos que, en ciertas condiciones, no tiene importancia este orden.

El primer ejemplo, y el más sencillo, se refiere a una «sucesión doble».

7.2 Ejemplo Para $m = 1, 2, 3, \ldots$ y $n = 1, 2, 3, \ldots$ sea

$$s_{m,n} = \frac{m}{m+n}.$$

Para todo n prefijado, será

$$\lim_{m \to \infty} s_{m,n} = 1,$$

de modo que

(4) $$\lim_{n \to \infty} \lim_{m \to \infty} s_{m,n} = 1.$$

Por otro lado, para todo m prefijado

$$\lim_{n \to \infty} s_{m,n} = 0,$$

de modo que

(5) $$\lim_{m\to\infty}\lim_{n\to\infty} s_{m,n} = 0.$$

7.3 Ejemplo Sea

$$f_n(x) = \frac{x^2}{(1+x^2)^n} \qquad (x \text{ real}; n = 0, 1, 2, \ldots),$$

y consideremos

(6) $$f(x) = \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} \frac{x^2}{(1+x^2)^n}.$$

Como $f_n(0) = 0$, tenemos que $f(0) = 0$. Para $x \neq 0$, la última serie de (6) es una serie geométrica convergente con suma $1 + x^2$ (Teorema 3.26). Por tanto,

(7) $$f(x) = \begin{cases} 0 & (x = 0), \\ 1 + x^2 & (x \neq 0), \end{cases}$$

de forma que una serie convergente de funciones continuas puede tener una suma discontinua.

7.4 Ejemplo Para $m = 1, 2, 3, \ldots$, hagamos

$$f_m(x) = \lim_{n\to\infty} (\cos m!\pi x)^{2n}.$$

Cuando $m!x$ es entero, $f_m(x) = 1$. Para todo otro valor de x, $f_m(x) = 0$. Sea ahora

$$f(x) = \lim_{m\to\infty} f_m(x).$$

Para x irracional, $f_m(x) = 0$ para todo m, por lo cual $f(x) = 0$. Para x racional, es decir, $x = p/q$, siendo p y q enteros, vemos que $m!x$ es entero si $m \geq q$, de modo que $f(x) = 1$. Por tanto

(8) $$\lim_{m\to\infty}\lim_{n\to\infty} (\cos m!\pi x)^{2n} = \begin{cases} 0 & (x \text{ irracional}), \\ 1 & (x \text{ racional}). \end{cases}$$

Hemos obtenido, así, una función límite discontinua en todas partes que no es integrable, según Riemann (Ejer. 4, Cap. 6).

7.5 Ejemplo Sea

(9) $$f_n(x) = \frac{\operatorname{sen} nx}{\sqrt{n}} \qquad (x \text{ real}, n = 1, 2, 3, \ldots),$$

y

$$f(x) = \lim_{n \to \infty} f_n(x) = 0.$$

Entonces, $f'(x) = 0$, y

$$f_n'(x) = \sqrt{n} \cos nx,$$

de modo que $\{f_n'\}$ no converge hacia f'. Por ejemplo,

$$f_n'(0) = \sqrt{n} \to +\infty$$

cuando $n \to \infty$, mientras que $f'(0) = 0$.

7.6 Ejemplo Sea

(10) $$f_n(x) = n^2 x(1 - x^2)^n \qquad (0 \le x \le 1, n = 1, 2, 3, \ldots).$$

Para $0 < x \le 1$, tenemos

$$\lim_{n \to \infty} f_n(x) = 0,$$

por el Teorema 3.20(*d*). Como $f_n(0) = 0$, vemos que

(11) $$\lim_{n \to \infty} f_n(x) = 0 \qquad (0 \le x \le 1).$$

Un cálculo sencillo demuestra que

$$\int_0^1 x(1 - x^2)^n \, dx = \frac{1}{2n + 2}.$$

Así, pues, a pesar de (11)

$$\int_0^1 f_n(x) \, dx = \frac{n^2}{2n + 2} \to +\infty$$

cuando $n \to \infty$.

Si sustituimos en (10) n^2 por n, se cumple todavía (11), pero, tenemos

$$\lim_{n \to \infty} \int_0^1 f_n(x) \, dx = \lim_{n \to \infty} \frac{n}{2n + 2} = \frac{1}{2},$$

cuando

$$\int_0^1 \left[\lim_{n \to \infty} f_n(x) \right] dx = 0.$$

Así, pues, el límite de la integral no es necesariamente igual a la integral del límite, aun cuando los dos sean finitos.

Después de estos ejemplos que demuestran que podemos cometer un error si se invierte el orden del proceso de límites descuidadamente, definiremos una nueva forma de convergencia, más severa que la puntual que aparece en la Definición 7.1, que nos permitirá llegar a resultados positivos.

CONVERGENCIA UNIFORME

7.7 Definición Decimos que una sucesión de funciones $\{f_n\}$ con $n = 1, 2, 3, \ldots$, converge *uniformemente* en E hacia una función f si para cada $\varepsilon > 0$ hay un entero N tal que $n \geq N$ implica

(12) $$|f_n(x) - f(x)| \leq \varepsilon$$

para todo $x \in E$.

Es claro que toda sucesión uniformemente convergente es puntualmente convergente. Concretamente, la diferencia entre los dos conceptos es la siguiente: si $\{f_n\}$ converge puntualmente en E, existe una función f tal que, para todo $\varepsilon > 0$ y cada $x \in E$, hay un entero N, que depende de ε y de x, tal que se cumple (12) si $n \geq N$; si $\{f_n\}$ converge uniformemente en E, es posible hallar, para cada $\varepsilon > 0$ *un* entero N tal que lo hace para *todo* $x \in E$.

Decimos que la serie $\Sigma f_n(x)$ converge uniformemente en E si la sucesión $\{s_n\}$ de sumas parciales definidas por

$$\sum_{i=1}^{n} f_i(x) = s_n(x)$$

converge uniformemente en E.

El criterio de Cauchy sobre convergencia uniforme es el siguiente:

7.8 Teorema *La sucesión de funciones* $\{f_n\}$*, definida en E, converge uniformemente en E si y solo si para cada* $\in > 0$ *existe un entero N tal que* $m \geq N$*;* $n \geq N$*;* $x \in E$ *implica*

(13) $$|f_n(x) - f_m(x)| \leq \varepsilon.$$

Demostración Supongamos que $\{f_n\}$ converge uniformemente en E, y sea f la función límite. Entonces, existe un entero N tal que $n \geq N$ y $x \in E$ implica

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{2},$$

de modo que

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \leq \varepsilon$$

si $n \geq N$; $m \geq N$, y $x \in E$.

Inversamente, supongamos que se cumple la condición de Cauchy. Por el Teorema 3.11, la sucesión $\{f_n(x)\}$ converge, para todo x, hacia un límite que podemos llamar $f(x)$. Así pues, la sucesión $\{f_n\}$ converge en E hacia f. Tenemos que demostrar que la convergencia es uniforme.

Sea $\varepsilon > 0$ dado, y elijamos N tal que se cumpla (13). Fijemos n y hagamos $m \to \infty$ en (13). Como $f_m(x) \to f(x)$ cuando $m \to \infty$, esto da

(14) $$|f_n(x) - f(x)| \leq \varepsilon$$

para cada $n \geq N$ y todo $x \in E$, lo que completa la demostración.

El criterio siguiente es útil algunas veces.

7.9 Teorema *Supongamos*

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad (x \in E).$$

Hagamos

$$M_n = \sup_{x \in E} |f_n(x) - f(x)|.$$

Entonces, $f_n \to f$ uniformemente en E si y solo si $M_n \to 0$ cuando $n \to \infty$.

Como esto es consecuencia inmediata de la Definición 7.7, omitimos los detalles de la demostración.

Hay un criterio de convergencia uniforme muy conveniente para las series, debido a Weierstrass:

7.10 Teorema *Supongamos que $\{f_n\}$ es una sucesión de funciones definidas en E, y supongamos*

$$|f_n(x)| \leq M_n \qquad (x \in E, n = 1, 2, 3, \ldots).$$

En estas condiciones, Σf_n converge uniformemente en E si ΣM_n converge.

Obsérvese que no se afirma la inversa (y de hecho, no es cierta).

Demostración Si ΣM_n converge, para $\varepsilon > 0$ arbitrario,

$$\left| \sum_{i=n}^{m} f_i(x) \right| \leq \sum_{i=n}^{m} M_i \leq \varepsilon \qquad (x \in E),$$

con tal que m y n sean suficientemente grandes. La convergencia uniforme, se deduce del Teorema 7.8.

CONVERGENCIA UNIFORME Y CONTINUIDAD

7.11 Teorema *Supongamos que $f_n \to f$ uniformemente en un conjunto E en un espacio métrico. Sea x un punto de acumulación de E, y supongamos que*

(15) $$\lim_{t \to x} f_n(t) = A_n \qquad (n = 1, 2, 3, \ldots).$$

$\{A_n\}$ convergerá, y

(16) $$\lim_{t \to x} f(t) = \lim_{n \to \infty} A_n.$$

En otras palabras, la conclusión es que

(17) $$\lim_{t \to x} \lim_{n \to \infty} f_n(t) = \lim_{n \to \infty} \lim_{t \to x} f_n(t).$$

Demostraci n Sea $\varepsilon > 0$ dado. Por la convergencia uniforme de $\{f_n\}$, existe N tal que $n \geq N$; $m \geq N$, y $t \in E$ implica

(18) $$|f_n(t) - f_m(t)| \leq \varepsilon.$$

Haciendo $t \to x$ en (18), obtenemos

$$|A_n - A_m| \leq \varepsilon$$

para $n \geq N$, $m \geq N$, de modo que $\{A_n\}$ es una sucesión de Cauchy y, por tanto, converge, digamos que hacia A.

Ahora,

(19) $$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|.$$

Elijamos, primeramente n de modo que

(20) $$|f(t) - f_n(t)| \leq \frac{\varepsilon}{3}$$

para todo $t \in E$ (lo que es posible por la convergencia uniforme) y tal que

(21) $$|A_n - A| \leq \frac{\varepsilon}{3}.$$

Entonces, para este n, elijamos una vecindad V de x tal que

$$|f_n(t) - A_n| \leq \frac{\varepsilon}{3} \tag{22}$$

Si $t \in V \cap E$, $t \neq x$.

Sustituyendo las desigualdades (20) a (22) en (19), vemos que

$$|f(t) - A| \leq \varepsilon,$$

siempre que $t \in V \cap E$, $t \neq x$. Lo que equivale a (16).

7.12 Teorema *Si $\{f_n\}$ es una sucesión de funciones continuas en E, y si $f_n \to f$ uniformemente en E, f es continua en E.*

Este importantísimo resultado es un corolario inmediato del Teorema 7.11.

El inverso no es cierto; esto es, una sucesión de funciones continuas puede converger en una función continua, aunque la convergencia no sea uniforme. En 7.6 se tiene un ejemplo de ello (para verlo, aplicar el Teorema 7.9). Pero hay casos en los que podemos afirmar el inverso:

7.13 Teorema *Si K es compacto, y*

(a) $\{f_n\}$ es una sucesión de funciones continuas sobre K,
(b) $\{f_n\}$ converge puntualmente a una función continua f sobre K,
(c) $f_n(x) \geq f_{n+1}(x)$ para todo $x \in K$, $n = 1, 2, 3, \ldots$,
Entonces $f_n \to f$ uniformemente sobre K.

Demostración Si se hace $g_n = f_n - f$. Entonces g_n es continua, $g_n \to 0$ puntualmente, y $g_n \geq g_{n+1}$. Se tiene que demostrar ahora que $g_n \to 0$ uniformemente sobre K.

Dado $\varepsilon > 0$, sea K_n el conjunto de todos los $x \in K$ con $g_n(x) \geq \varepsilon$. Como g_n es continua, K_n es cerrado (véase el Teorema 4.8), por ende también es compacto (véase el Teorema 2.35). Ya que $g_n \geq g_{n+1}$, se tiene $K_n \supset K_{n+1}$. Fijemos $x \in K$. Debido a que $g_n(x) \to 0$, es evidente que $x \notin K_n$ si n es suficientemente grande. Por esto $x \notin \bigcap K_n$. Dicho de otro modo, $\bigcap K_n$ es vacío.

Por lo tanto K_N es vacío para algún N (véase el Teorema 2.36). Con esto se deduce que $0 \leq g_n(x) < \varepsilon$ para todo $x \in K$ y todo $n \geq N$. Esto demuestra el teorema.

Nótese que la compactibilidad es necesaria aquí. Por ejemplo, si

$$f_n(x) = \frac{1}{nx + 1} \qquad (0 < x < 1;\ n = 1, 2, 3, \ldots)$$

entonces $f_n(x) \to 0$ monótonamente en $(0, 1)$, pero la convergencia no es uniforme.

7.14 Definición Si X es un espacio métrico, $\mathscr{C}(X)$ representará al conjunto de todas las funciones valuadas en los complejos con dominio X, continuas y acotadas.

[Nótese que la acotabilidad es redundante si X es compacto (véase el Teorema 4.15). Entonces $\mathscr{C}(X)$ consta de todas las funciones complejas continuas definidas sobre X si X es compacto.]

A cada $f \in \mathscr{C}(X)$ se le asocia su *norma suprema*

$$\|f\| = \sup_{x \in X} |f(x)|.$$

Como se admite que f es acotada, $\|f\| < \infty$. Es obvio que $\|f\| = 0$ solo si $f(x) = 0$ para cada $x \in X$, es decir, solo si $f = 0$. Si $h = f + g$, entonces

$$|h(x)| \leq |f(x)| + |g(x)| \leq \|f\| + \|g\|$$

para todo $x \in X$; de aquí que

$$\|f+g\| \leq \|f\| + \|g\|.$$

Si se define la distancia entre $f \in \mathscr{C}(X)$ y $g \in \mathscr{C}(X)$ como $\|f - g\|$, se deduce que se cumplen los axiomas 2.15 para una métrica.

Entonces se ha vuelto $\mathscr{C}(X)$ un espacio métrico.

El Teorema 7.9 puede volver a redactarse como sigue:

Una sucesión $\{f_n\}$ converge hacia f con respecto a la métrica de $\mathscr{C}(X)$ si, y solo si $f_n \to f$ uniformemente sobre X.

De acuerdo con esto, los subconjuntos cerrados de $\mathscr{C}(X)$ se llaman a veces conjuntos *uniformemente cerrados*, la cerradura de un conjunto $\mathscr{A} \subset \mathscr{C}(X)$ se dice que es *cerradura uniforme*, y así sucesivamente.

7.15 Teorema *La métrica anterior hace de $\mathscr{C}(X)$ un espacio métrico completo.*

Demostración Sea $\{f_n\}$ una sucesión de Cauchy en $\mathscr{C}(X)$. Esto quiere decir que a cada $\varepsilon > 0$ le corresponde un N tal que $\|f_n - f_m\| < \varepsilon$ si $n \geq N$ y $m \geq N$. Se deduce (por el Teorema 7.8) que hay una función f con dominio X para la cual $\{f_n\}$ converge uniformemente. Y por el Teorema 7.12, f es continua. Además, f es acotada, porque hay un n tal que $|f(x) - f_n(x)| < 1$ para todo $x \in X$, y f_n es acotada.

Entonces $f \in \mathscr{C}(X)$, y debido a que $f_n \to f$ uniformemente sobre X, se tiene que $\|f - f_n\| \to 0$ cuando $n \to \infty$.

CONVERGENCIA UNIFORME E INTEGRACIÓN

7.16 Teorema *Sea α monótona creciente sobre $[a,b]$. Supóngase que $f_n \in \mathscr{R}(\alpha)$ sobre $[a,b]$, para $n = 1, 2, 3, \ldots$, y que $f_n \to f$ uniformemente sobre $[a,b]$. Entonces $f \in \mathscr{R}(\alpha)$ sobre $[a,b]$, y*

$$\int_a^b f\,d\alpha = \lim_{n\to\infty} \int_a^b f_n\,d\alpha. \tag{23}$$

(La existencia del límite en (23) es parte de la conclusión.)

Demostración Es suficiente demostrarlo para f_n real. Se hace

$$\varepsilon_n = \sup |f_n(x) - f(x)|, \tag{24}$$

el supremum se ha tomado sobre $a \le x \le b$. Entonces

$$f_n - \varepsilon_n \le f \le f_n + \varepsilon_n,$$

de manera que las integrales superior e inferior de f satisfacen (véase la Definición 6.2),

$$\int_a^b (f_n - \varepsilon_n)\,d\alpha \le \underline{\int} f\,d\alpha \le \overline{\int} f\,d\alpha \le \int_a^b (f_n + \varepsilon_n)\,d\alpha. \tag{25}$$

En consecuencia

$$0 \le \overline{\int} f\,d\alpha - \underline{\int} f\,d\alpha \le 2\varepsilon_n[\alpha(b) - \alpha(a)].$$

Como $\varepsilon_n \to 0$ cuando $n \to \infty$ (por el Teorema 7.9), las integrales superior e inferior de f son iguales.

Entonces $f \in \mathscr{R}(\alpha)$. Otra aplicación de (25) produce ahora

$$\left|\int_a^b f\,d\alpha - \int_a^b f_n\,d\alpha\right| \le \varepsilon_n[\alpha(b) - \alpha(a)]. \tag{26}$$

Esto implica (23).

Corolario Si $f_n \in \mathscr{R}(\alpha)$ *en* $[a,b]$, *y si*

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \qquad (a \le x \le b),$$

convergiendo la serie uniformemente en $[a,b]$, *entonces*

$$\int_a^b f\,d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n\,d\alpha.$$

En otras palabras, las series pueden ser integradas término a término.

CONVERGENCIA UNIFORME Y DIFERENCIACIÓN

Hemos visto ya, en el Ejemplo 7.5, que la convergencia uniforme de $\{f_n\}$ no implica nada sobre la sucesión $\{f_n'\}$. Así pues, se necesitan hipótesis más rigurosas para asegurar que $f_n' \to f'$ si $f_n \to f$.

7.17 Teorema *Supongamos que* $\{f_n\}$ *es una sucesión de funciones, diferenciables en* $[a,b]$ *de modo que* $\{f_n(x_0)\}$ *converge para algún punto* x_0 *en* $[a,b]$. *Si* $\{f_n'\}$ *converge uniformemente en* $[a,b]$, $\{f_n\}$ *converge uniformemente en* $[a,b]$ *hacia una función* f, *y*

(27) $$f'(x) = \lim_{n\to\infty} f_n'(x) \qquad (a \le x \le b).$$

Demostración Sea $\varepsilon > 0$ dado. Elijamos N tal que $n \ge N$, $m \ge N$, implica

(28) $$|f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2}$$

y

(29) $$|f_n'(t) - f_m'(t)| < \frac{\varepsilon}{2(b-a)} \qquad (a \le t \le b).$$

Si aplicamos el teorema del valor medio 5.19 a la función $f_n - f_m$, (29) demuestra que,

(30) $$|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \le \frac{|x-t|\varepsilon}{2(b-a)} \le \frac{\varepsilon}{2}$$

para cualesquiera x y t en $[a,b]$, si $n \ge N$, $m \ge N$. La desigualdad

$$|f_n(x) - f_m(x)| \le |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)|$$

implica, por (28) y (30), que

$$|f_n(x) - f_m(x)| < \varepsilon \qquad (a \le x \le b, n \ge N, m \ge N),$$

de modo que $\{f_n\}$ converge uniformemente en $[a,b]$. Sea

$$f(x) = \lim_{n\to\infty} f_n(x) \qquad (a \leq x \leq b).$$

Fijemos un punto x en $[a,b,]$ y definamos

(31) $$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}, \qquad \phi(t) = \frac{f(t) - f(x)}{t - x}$$

para $a \leq t \leq b$ y $t \neq x$. Será

(32) $$\lim_{t\to x} \phi_n(t) = f_n'(x) \qquad (n = 1, 2, 3, \ldots).$$

La primera desigualdad de (30) demuestra que

$$|\phi_n(t) - \phi_m(t)| \leq \frac{\varepsilon}{2(b-a)} \qquad (n \geq N, m \geq N),$$

de modo que $\{\phi_n\}$ converge uniformemente, para $t \neq x$. Como $\{f_n\}$ converge hacia f, deducimos de (31) que

(33) $$\lim_{n\to\infty} \phi_n(t) = \phi(t)$$

uniformemente para $a \leq t \leq b$ y $t \neq x$.

Si aplicamos, ahora, el Teorema 7.11 a $\{\phi_n\}$, (32) y (33) demuestra que

$$\lim_{t\to x} \phi(t) = \lim_{n\to\infty} f_n'(x);$$

y esta es (27), por la definición de $\phi(t)$.

Observación: Si además de las hipótesis anteriores, se admite la continuidad de las funciones f_n', se puede dar una demostración mucho más corta de (27), basada en el Teorema 7.16 y el teorema fundamental del cálculo.

7.18 Teorema *Existe una función real continua sobre la recta real que es no diferenciable en ninguna parte.*

Demostración Se define

(34) $$\varphi(x) = |x| \qquad (-1 \leq x \leq 1)$$

y se amplía la definición de $\varphi(x)$ para toda x real pidiendo que

(35) $$\varphi(x + 2) = \varphi(x).$$

Entonces, para todo s y t,

(36) $$|\varphi(s) - \varphi(t)| \leq |s - t|.$$

En particular, φ es continua sobre R^1. Se define

(37) $$f(x) = \sum_{n=0}^{\infty} (\tfrac{3}{4})^n \varphi(4^n x).$$

El Teorema 7.10 muestra que la serie (37) converge uniformemente sobre R^1, ya que $0 \leq \varphi \leq 1$. Por el Teorema 7.12, f es continua sobre R^1.

Si ahora se fijan un número real x y un entero positivo m y se hace

(38) $$\delta_m = \pm \tfrac{1}{2} \cdot 4^{-m}$$

donde el signo se elige de tal forma que no esté ningún entero entre $4^m x$ y $4^m(x + \delta_m)$. Esto puede hacerse ya que $4^m |\delta_m| = \frac{1}{2}$. Entonces se define

(39) $$\gamma_n = \frac{\varphi(4^n(x + \delta_m)) - \varphi(4^n x)}{\delta_m}.$$

Cuando $n > m$, *entonces* $4^n \delta_m$ es un entero par, así que $\gamma_n = 0$. Cuando $0 \leq n \leq m$, (36) implica que $|\gamma_n| \leq 4^n$.

Debido a que $|\gamma_m| = 4^m$, se concluye

$$\left|\frac{f(x + \delta_m) - f(x)}{\delta_m}\right| = \left|\sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \gamma_n\right| \geq 3^m - \sum_{n=0}^{m-1} 3^n = \tfrac{1}{2}(3^m + 1).$$

Conforme $m \to \infty$, $\delta_m \to 0$. Se deduce que f no es diferenciable en x.

FAMILIAS EQUICONTINUAS DE FUNCIONES

En el Teorema 3.6 vimos que toda sucesión acotada de números complejos contiene una subsucesión convergente; y se plantea, ahora, la pregunta de si sucede algo similar con las sucesiones de funciones. Para concretar más, definiremos dos clases de acotación.

7.19 Definición Sea $\{f_n\}$ una sucesión de funciones definidas en un conjunto E.

Decimos que $\{f_n\}$ es *acotada puntualmente* en E si la sucesión $\{f_n(x)\}$ es acotada para cada $x \in E$, esto es, si existe una función ϕ con valores finitos en E tal que

$$|f_n(x)| < \phi(x) \qquad (x \in E, n = 1, 2, 3, \ldots).$$

Decimos que $\{f_n\}$ es *uniformemente acotada* en E si existe un número M tal que

$$|f_n(x)| < M \qquad (x \in E, n = 1, 2, 3, \ldots).$$

Si $\{f_n\}$ es acotada puntualmente en E y E_1 es un subconjunto numerable de E, es siempre posible hallar una subsucesión $\{f_{n_k}\}$ tal que $\{f_{n_k}(x)\}$ converge para todo $x \in E_1$, lo que puede hacerse por el proceso de la diagonal, utilizado en la demostración del Teorema 7.23.

Sin embargo, aún si $\{f_n\}$ es una sucesión uniformemente acotada de funciones continuas en un conjunto compacto E, no es necesaria la existencia de una subsucesión que converja puntualmente en E. En el ejemplo dado a continuación, sería fatigoso demostrarlo con los medios de que disponemos hasta este momento, pero la demostración es muy sencilla si se recurre a un teorema del capítulo 11.

7.20 Ejemplo Sea

$$f_n(x) = \operatorname{sen} nx \qquad (0 \leq x \leq 2\pi, n = 1, 2, 3, \ldots).$$

Supongamos que existe una sucesión $\{n_k\}$ tal que $\{\operatorname{sen} n_k x\}$ converge para todo $x \in [0, 2\pi]$. En este caso, debemos tener

$$\lim_{k\to\infty} (\operatorname{sen} n_k x - \operatorname{sen} n_{k+1} x) = 0 \qquad (0 \leq x \leq 2\pi);$$

y de aquí

$$(40) \qquad \lim_{k\to\infty} (\operatorname{sen} n_k x - \operatorname{sen} n_{k+1} x)^2 = 0 \qquad (0 \leq x \leq 2\pi).$$

Por el teorema de Lebesgue, referente a la integración de sucesiones convergentes acotadas (Teorema 11.32), (40) implica

$$(41) \qquad \lim_{k\to\infty} \int_0^{2\pi} (\operatorname{sen} n_k x - \operatorname{sen} n_{k+1} x)^2 \, dx = 0.$$

Pero un cálculo sencillo muestra que

$$\int_0^{2\pi} (\operatorname{sen} n_k x - \operatorname{sen} n_{k+1} x)^2 \, dx = 2\pi,$$

lo que contradice a (41).

Otra duda que puede plantearse, es si toda sucesión convergente contiene una subsucesión uniformemente convergente. El próximo ejemplo demuestra que no sucede necesariamente así, aun si la sucesión es uniformemente acotada en un conjunto compacto. (El Ejemplo 7.6 demuestra que una sucesión de funciones acotadas puede converger sin ser uniformemente acotada; pero se ve inmediatamente que la convergencia uniforme de una sucesión de funciones acotadas implica la acotación uniforme.)

7.21 Ejemplo Sea

$$f_n(x) = \frac{x^2}{x^2 + (1 - nx)^2} \qquad (0 \le x \le 1, n = 1, 2, 3, \ldots).$$

Entonces, $|f_n(x)| \le 1$, de modo que $\{f_n\}$ es uniformemente acotada en $[0, 1]$. Además,

$$\lim_{n \to \infty} f_n(x) = 0 \qquad (0 \le x \le 1),$$

pero

$$f_n\left(\frac{1}{n}\right) = 1 \qquad (n = 1, 2, 3, \ldots),$$

de modo que ninguna subsucesión puede converger uniformemente en $[0, 1]$.

En la siguiente definición, se da el concepto de equicontinuidad que nos es necesario en este punto.

7.22 Definición Se dice que una familia $\mathscr{F}$ de funciones complejas f definidas en un conjunto E en un espacio métrico X es *equicontinua* en E, si para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que

$$|f(x) - f(y)| < \varepsilon$$

siempre que $d(x,y) < \delta$; $x \in E$; $y \in E$, y $f \in \mathscr{F}$. Aquí, d expresa la métrica de X.

Es claro que todo miembro de una familia equicontinua es uniformemente continua.

La sucesión del Ejemplo 7.21 no es equicontinua.

Los Teoremas 7.24 y 7.25 mostrarán que hay una relación bastante marcada entre la equicontinuidad y la convergencia uniforme de sucesiones de funciones continuas. Pero primero se describirá un proceso de selección que nada tiene que ver con la continuidad.

7.23 Teorema *Si $\{f_n\}$ es una sucesión puntualmente acotada de funciones complejas sobre un conjunto numerable E, entonces $\{f_n\}$ tiene una subsucesión $\{f_{n_k}\}$ tal que $\{f_{n_k}(x)\}$ converge para cada $x \in E$.*

Demostración Sean $\{x_i\}$, $i = 1, 2, 3, \ldots$, los puntos de E, ordenados en una sucesión. Como $\{f_n(x_1)\}$ es acotada, existe una subsucesión que se representará por $\{f_{1,k}\}$, tal que $\{f_{1,k}(x_1)\}$ converge cuando $k \to \infty$.

Considérense ahora las sucesiones S_1, S_2, $S_3, \ldots$, que se representarán por medio del arreglo

$$\begin{array}{llllll} S_1: & f_{1,1} & f_{1,2} & f_{1,3} & f_{1,4} & \cdots \\ S_2: & f_{2,1} & f_{2,2} & f_{2,3} & f_{2,4} & \cdots \\ S_3: & f_{3,1} & f_{3,2} & f_{3,3} & f_{3,4} & \cdots \end{array}$$

y que tienen la siguientes propiedades:

(*a*) S_n es una subsucesión de S_{n-1}, para $n = 2, 3, 4, \ldots$.
(*b*) $\{f_{n,k}(x_n)\}$ converge, cuando $k \to \infty$ (la acotabilidad de $\{f_n(x_n)\}$ hace posible que S_n pueda elegirse de esta forma);
(*c*) El orden en el que aparecen las funciones es el mismo en cada sucesión; es decir, si una función precede a otra en S_1, están en la misma relación en cada S_n, hasta que desaparecen una u otra. De aquí que, cuando se pasa una fila del cuadro anterior a la inmediatamente inferior, las funciones pueden transladarse hacia la izquierda, pero nunca hacia la derecha.

Sigamos ahora, la diagonal del arreglo; es decir, considérese la sucesión

$$S: \quad f_{1,1} \quad f_{2,2} \quad f_{3,3} \quad f_{4,4} \cdots.$$

De (*c*) se tiene que la sucesión S (excepto posiblemente sus primeros $n - 1$ términos) es una subsucesión de S_n, para $n = 1, 2, 3, \ldots$. En consecuencia (*b*) implica que $\{f_{n,n}(x_i)\}$ converge, cuando $n \to \infty$, para cada $x_i \in E$.

7.24 Teorema *Si K es un espacio métrico compacto y $f_n \in \mathscr{C}(K)$ para $n = 1, 2, 3, \ldots$, y si $\{f_n\}$ converge uniformemente sobre K, entonces $\{f_n\}$ es equicontinua sobre K.*

Demostración Sea $\varepsilon > 0$ dado. Como $\{f_n\}$ converge uniformemente, hay un entero N tal que

$$\|f_n - f_N\| < \varepsilon \qquad (n > N). \tag{42}$$

(Véase la Definición 7.14.) Debido a que las funciones continuas son uniformemente continuas sobre conjuntos compactos, hay un $\delta > 0$ tal que

(43) $$|f_i(x) - f_i(y)| < \varepsilon$$

si $1 \le i \le N$ y $d(x,y) < \delta$.

Si $n > N$ y $d(x,y)$, se deduce que

$$|f_n(x) - f_n(y)| \le |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < 3\varepsilon.$$

Y junto con (43), queda demostrado el teorema.

7.25 Teorema *si K es compacto, $f_n \in \mathscr{C}(K)$ para $n = 1, 2, 3, \ldots,$ y si $\{f_n\}$ es acotada puntualmente y equicontinua sobre K, entonces*

(*a*) *$\{f_n\}$ es uniformemente acotada sobre K,*
(*b*) *$\{f_n\}$ contiene una subsucesión uniformemente convergente.*

Demostración

(*a*) Sea $\varepsilon > 0$ conocido y elíjase $\delta > 0$, de acuerdo con la Definición 7.22, de manera que

(44) $$|f_n(x) - f_n(y)| < \varepsilon$$

para todo n, siempre que $d(x,y) < \delta$.

Como K es compacto, hay un número finito de puntos $p_1, \ldots, p_r$ en K tal que para cada $x \in K$ corresponde al menos un p_i con $d(x,p_i) < \delta$. Ya que $\{f_n\}$ es acotada puntualmente, existe $M_i < \infty$ tal que $|f_n(p_i)| < M_i$ para todo n. Si $M = \text{máx}\,(M_1, \ldots, M_r)$, entonces $|f(x)| < M + \varepsilon$ para cada $x \in K$. Esto demuestra (*a*).

(*b*) Sea E un subconjunto denso numerable de K. (Véase el Ejercicio 25 del Cap. 2 para la existencia de tal conjunto E.) El Teorema 7.23 muestra que $\{f_n\}$ tiene una subsucesión $\{f_{n_i}\}$ tal que $\{f_{n_i}(x)\}$ converge para cada $x \in E$.

Haciendo $f_{n_i} = g_i$, con el propósito de simplificar la notación. Se probará que $\{g_i\}$ converge uniformemente sobre K.

Sea $\varepsilon > 0$, y tómese $\delta > 0$ como al inicio de esta demostración. Sea $V(x,\delta)$ el conjunto de todos los $y \in K$ con $d(x,y) < \delta$. Como E es denso en K, y K es compacto, hay un número finito de puntos $x_1, \ldots, x_m$ en E tales que

(45) $$K \subset V(x_1, \delta) \cup \cdots \cup V(x_m, \delta).$$

Ya que $\{g_i(x)\}$ converge para cada $x \in E$, hay un entero N tal que

(46) $$|g_i(x_s) - g_j(x_s)| < \varepsilon$$

siempre que $i \geq N, j \geq N$, $1 \leq s \leq m$.

Si $x \in K$, (45) muestra que $x \in V(x_s, \delta)$ para algún s, de manera que

$$|g_i(x) - g_i(x_s)| < \varepsilon$$

para cada i. Si $i \geq N$ y $j \geq N$, se deduce de (46) que

$$\begin{aligned}|g_i(x) - g_j(x)| &\leq |g_i(x) - g_i(x_s)| + |g_i(x_s) - g_j(x_s)| + |g_j(x_s) - g_j(x)| \\ &< 3\varepsilon.\end{aligned}$$

Esto completa la demostración.

TEOREMA DE STONE-WEIERSTRASS

7.26 Teorema *Si f es una función compleja continua en $[a,b]$, existe una sucesión de polinomios P_n, tal que*

$$\lim_{n\to\infty} P_n(x) = f(x)$$

uniformemente en $[a,b,]$. Si f es real, se deben tomar los P_n reales.

Esta es la forma en que el teorema fue enunciado originalmente por Weierstrass.

Demostración Admitiremos, sin pérdida de generalidad, que $[a,b,] = [0, 1]$, y que $f(0) = f(1) = 0$. puesto que si se demuestra el teorema para este caso, consideraremos

$$g(x) = f(x) - f(0) - x[f(1) - f(0)] \qquad (0 \leq x \leq 1).$$

Aquí $g(0) = g(1) = 0$, y si puede obtenerse g como límite de una sucesión uniformemente convergente de polinomios, es claro que lo mismo es cierto para f, pues, $f - g$ es un polinomio.

Además, supondremos que $f(x)$ es cero para x fuera de $[0, 1]$. Entonces, f es uniformemente continua en toda la recta.

Hacemos

(47) $$Q_n(x) = c_n(1 - x^2)^n \qquad (n = 1, 2, 3, \ldots),$$

donde se ha elegido c_n, de modo que

(48) $$\int_{-1}^{1} Q_n(x)\,dx = 1 \qquad (n = 1, 2, 3, \ldots).$$

Necesitamos algún conocimiento sobre el orden de magnitud de c_n. Como

$$\begin{aligned}\int_{-1}^{1} (1-x^2)^n\,dx = 2\int_{0}^{1} (1-x^2)^n\,dx &\geq 2\int_{0}^{1/\sqrt{n}} (1-x^2)^n\,dx \\ &\geq 2\int_{0}^{1/\sqrt{n}} (1-nx^2)\,dx \\ &= \frac{4}{3\sqrt{n}} \\ &> \frac{1}{\sqrt{n}},\end{aligned}$$

de (48) se deduce que

(49) $$c_n < \sqrt{n}.$$

La desigualdad $(1 - x^2)^n \geq 1 - nx^2$ que se usó antes, se comprueba fácilmente que es cierta, considerando la función

$$(1-x^2)^n - 1 + nx^2$$

que es cero en $x = 0$, y cuya derivada es positiva en $(0, 1)$.

Para todo $\delta > 0$ (49) implica

(50) $$Q_n(x) \leq \sqrt{n}\,(1-\delta^2)^n \qquad (\delta \leq |x| \leq 1),$$

de modo que $Q_n \to 0$ uniformemente en $\delta \leq |x| \leq 1$.

Sea, ahora,

(51) $$P_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)\,dt \qquad (0 \leq x \leq 1).$$

Nuestras hipótesis sobre f demuestran, por un simple cambio de variable, que

$$P_n(x) = \int_{-x}^{1-x} f(x+t)Q_n(t)\,dt = \int_{0}^{1} f(t)Q_n(t-x)\,dt,$$

y la última integral es un polinomio en x. Así, pues, $\{P_n\}$ es una sucesión de polinomios, que son reales si f es real.

Dado $\varepsilon > 0$, elegiremos $\delta > 0$, tal que $|y - x| < \delta$ implica

$$|f(y) - f(x)| < \frac{\varepsilon}{2}.$$

Sea $M = \sup |f(x)|$. Utilizando (48), (50) y el hecho de ser $Q_n(x) \geq 0$, vemos que para $0 \leq x \leq 1$.

$$\begin{aligned}
|P_n(x) - f(x)| &= \left| \int_{-1}^{1} [f(x+t) - f(x)] Q_n(t)\, dt \right| \\
&\leq \int_{-1}^{1} |f(x+t) - f(x)| Q_n(t)\, dt \\
&\leq 2M \int_{-1}^{-\delta} Q_n(t)\, dt + \frac{\varepsilon}{2} \int_{-\delta}^{\delta} Q_n(t)\, dt + 2M \int_{\delta}^{1} Q_n(t)\, dt \\
&\leq 4M\sqrt{n}\,(1 - \delta^2)^n + \frac{\varepsilon}{2} \\
&< \varepsilon
\end{aligned}$$

para todo n suficientemente grande, lo que demuestra el teorema.

Es instructivo representar las gráficas de Q_n para algunos valores de n; nótese también, que necesitamos la continuidad uniforme de f para deducir la convergencia uniforme de $\{P_n\}$.

En la demostración del Teorema 7.32, no necesitamos todo el rigor del Teorema 7.26, sino solamente el siguiente caso particular, que enunciamos como corolario.

7.27 Corolario *Para cada intervalo $[-a,a]$ hay una sucesión de polinomios reales P_n, tal que $P_n(0) =$ y que*

$$\lim_{n \to \infty} P_n(x) = |x|$$

uniformemente en $[-a,a]$.

Demostración Por el Teorema 7.26, existe una sucesión $\{P_n^*\}$ de polinomios reales que converge hacia $|x|$ uniformemente en $[-a,a]$. En particular, $P_n^*(0) \to 0$ cuando $n \to \infty$. Los polinomios

$$P_n(x) = P_n^*(x) - P_n^*(0) \qquad (n = 1, 2, 3, \ldots)$$

tienen las propiedades deseadas.

Ahora, aislaremos las propiedades de los polinomios que hacen posible el teorema de Weierstrass.

7.28 Definición Se dice que una familia $\mathscr{A}$ de funciones complejas definidas en un conjunto E es un *álgebra* si (i) $f + g \in \mathscr{A}$ (ii) $fg \in \mathscr{A}$, y (iii) $cf \in \mathscr{A}$ para toda $f \in \mathscr{A}$, $g \in \mathscr{A}$ y para todas las constantes complejas c, esto es, si $\mathscr{A}$ es cerrada, respecto de la adición, multiplicación y multiplicación escalar. Tendremos que considerar también álgebras de funciones reales; en este caso, es natural que solo debe cumplirse (iii) para todo número real c.

Si $\mathscr{A}$ tiene la propiedad de ser $f \in \mathscr{A}$ cuando $f_n \in \mathscr{A}$ $(n = 1, 2, 3, \ldots)$ y $f_n \to f$ uniformemente en E, se dice que $\mathscr{A}$ es *uniformemente cerrada.*

Sea $\mathscr{B}$ el conjunto de todas las funciones que son límite de sucesiones uniformemente convergentes de elementos de $\mathscr{A}$. A $\mathscr{B}$ se le llama *cerradura uniforme* de $\mathscr{A}$.

Por ejemplo, el conjunto de todos los polinomios es un álgebra, y puede enunciarse el teorema de Weierstrass diciendo que el conjunto de funciones continuas en $[a,b]$ es la cerradura uniforme del conjunto de polinomios en $[a,b]$.

7.29 Teorema *Sea $\mathscr{B}$ la cerradura uniforme de un álgebra $\mathscr{A}$ de funciones acotadas. $\mathscr{B}$ es un álgebra uniformemente cerrada.*

Demostración Si $f \in \mathscr{B}$ y $g \in \mathscr{B}$, existen sucesiones uniformemente convergentes $\{f_n\}$, $\{g_n\}$, tales que $f_n \to f$, $g_n \to g$ y $f_n \in \mathscr{A}$, $g_n \in \mathscr{A}$. Como estamos tratando con funciones acotadas, es fácil demostrar que

$$f_n + g_n \to f + g, \qquad f_n g_n \to fg, \qquad cf_n \to cf,$$

siendo c una constante cualquiera, y la convergencia uniforme en cada caso.

De aquí, que $f + g \in \mathscr{B}$; $fg \in \mathscr{B}$, y $cf \in \mathscr{B}$, de modo que $\mathscr{B}$ es un álgebra.

Por el Teorema 2.27, $\mathscr{B}$ es cerrada (uniformemente).

7.30 Definición Sea $\mathscr{A}$ una familia de funciones en un conjunto E. Se dice que $\mathscr{A}$ *separa puntos* en E si a cada par de puntos distintos x_1, $x_2 \in E$ corresponde una función $f \in \mathscr{A}$, tal que $f(x_1) \neq f(x_2)$.

Si a cada $x \in E$ corresponde una función $g \in \mathscr{A}$, tal que $g(x) \neq 0$, decimos que $\mathscr{A}$ *no desaparece en ningún punto de E.*

El álgebra de todos los polinomios en una variable, tiene, como se ve fácilmente, esas propiedades en R^1. Un ejemplo de un álgebra que no separa puntos, es el conjunto de todos los polinomios pares, por ejemplo, en $[-1, 1]$, pues, $f(-x) = f(x)$ para toda función par f.

El teorema siguiente aclara más estos conceptos.

7.31 Teorema *Supongamos que $\mathscr{A}$ es un álgebra de funciones en un conjunto E, que $\mathscr{A}$ separa puntos en E, y no desaparece en ningún punto de E.*

Sean x_1, x_2 puntos distintos de E, y c_1, c_2 constantes (reales si $\mathscr{A}$ es un álgebra real). Entonces, $\mathscr{A}$ contiene una función f, tal que

$$f(x_1) = c_1, \qquad f(x_2) = c_2.$$

Demostración *Las suposiciones muestran que $\mathscr{A}$ contiene funciones g, h* y *k* tales que

$$g(x_1) \neq g(x_2), \qquad h(x_1) \neq 0, \qquad k(x_2) \neq 0.$$

Haciendo

$$u = gk - g(x_1)k, \qquad v = gh - g(x_2)h.$$

Entonces $u \in \mathscr{A}$, $v \in \mathscr{A}$, $u(x_1) = v(x_2) = 0$, $u(x_2) \neq 0$, y $v(x_1) \neq 0$. Por tanto

$$f = \frac{c_1 v}{v(x_1)} + \frac{c_2 u}{u(x_2)}$$

tiene las propiedades requeridas.

Se tiene ya todo el material necesario para la generalización del teorema de Weierstrass, debida a Stone.

7.32 Teorema *Sea $\mathscr{A}$ un álgebra de funciones reales continuas sobre un conjunto compacto K. Si $\mathscr{A}$ separa puntos sobre K y si $\mathscr{A}$ no desaparece en ningún punto de K, entonces la cerradura uniforme $\mathscr{B}$ de $\mathscr{A}$ consta de todas las funciones reales continuas sobre K.*

Se hará la demostración en varias etapas, cuatro para ser exactos.

1ª ETAPA *Si $f \in \mathscr{B}$, entonces $|f| \in \mathscr{B}$.*

Demostración Sea

$$a = \sup |f(x)| \qquad (x \in K) \tag{52}$$

y $\varepsilon > 0$ conocidos. Por el Corolario 7.27 existen números reales $c_1, \ldots, c_n$ tales que

$$\left| \sum_{i=1}^{n} c_i y^i - |y| \right| < \varepsilon \qquad (-a \leq y \leq a). \tag{53}$$

Como $\mathscr{B}$ es un álgebra, la función

$$g = \sum_{i=1}^{n} c_i f^i$$

es un miembro de $\mathscr{B}$. De (52) y (53) se tiene que

$$\big| g(x) - |f(x)| \big| < \varepsilon \qquad (x \in K).$$

Debido a que $\mathscr{B}$ es uniformemente cerrada, es evidente que $|f| \in \mathscr{B}$. ∈

2ª ETAPA *Si $f \in \mathscr{B}$ y $g \in \mathscr{B}$, entonces* máx $(f,g) \in \mathscr{B}$ *y* mín $(f,g) \in \mathscr{B}$.

Se debe entender por máx (f,g), la función h definida como

$$h(x) = \begin{cases} f(x) & \text{si } f(x) \geq g(x), \\ g(x) & \text{si } f(x) < g(x), \end{cases}$$

y mín (f,g) se define de la misma manera.

Demostración La 2ª etapa se deduce a partir de la 1ª y de las identidades siguientes:

$$\text{máx}\,(f,g) = \frac{f+g}{2} + \frac{|f-g|}{2},$$

$$\text{mín}\,(f,g) = \frac{f+g}{2} - \frac{|f-g|}{2}.$$

Por iteración, es posible extender el resultado para un conjunto finito cualquiera de funciones: Si $f_1, \ldots, f_n \in \mathscr{B}$, entonces máx $(f_1, \ldots, f_n) \in \mathscr{B}$, y

$$\text{mín}\,(f_1, \ldots, f_n) \in \mathscr{B}.$$

3ª ETAPA *Dados una función real continua f sobre K, un punto $x \in K$, y $\varepsilon > 0$, existe entonces una función $g_x \in \mathscr{B}$* tal que $g_x(x) = f(x)$ *y*

(54) $$g_x(t) > f(t) - \varepsilon \qquad (t \in K).$$

Demostración Como $\mathscr{A} \subset \mathscr{B}$ y $\mathscr{A}$ satisfacen las hipótesis del Teorema 7.31, también lo hace $\mathscr{B}$. Por consiguiente, para cada $y \in K$, se puede encontrar una función $h_y \in \mathscr{B}$ tal que

(55) $$h_y(x) = f(x), \qquad h_y(y) = f(y).$$

Como h_y es continua, existe un conjunto abierto J_y que contiene a y tal que

(56) $$h_y(t) > f(t) - \varepsilon \qquad (t \in J_y).$$

Debido a que K es compacto, existe un conjunto finito de puntos $y_1, \ldots, y_n$ tal que

(57) $$K \subset J_{y_1} \cup \cdots \cup J_{y_n}.$$

Haciendo

$$g_x = \text{máx}\,(h_{y_1}, \ldots, h_{y_n}).$$

Por la 2ª etapa, $g \in \mathscr{B}$ y las relaciones (55) a (57) muestran que g_x tiene las otras propiedades requeridas.

4ª ETAPA *Dados, una función real continua f sobre K, y $\varepsilon > 0$, existe una función $h \in \mathscr{B}$ tal que*

(58) $$|h(x) - f(x)| < \varepsilon \qquad (x \in K).$$

Como $\mathscr{B}$ es uniformemente cerrada, esta proposición es equivalente a la conclusión del teorema.

Demostración Considérense las funciones g_x, para cada $x \in K$, que se construyeron en la 3ª etapa. Por la continuidad de g_x, existen conjuntos abiertos V_x que contienen a x, tales que

(59) $$g_x(t) < f(t) + \varepsilon \qquad (t \in V_x).$$

Como K es compacto, existe un conjunto finito de puntos $x_1, \ldots, x_m$ tales que

(60) $$K \subset V_{x_1} \cup \cdots \cup V_{x_m}.$$

Haciendo

$$h = \text{mín}\,(g_{x_1}, \ldots, g_{x_m}).$$

Por la 2ª etapa se tiene que $h \in \mathscr{B}$, y (54) implica que

(61) $$h(t) > f(t) - \varepsilon \qquad (t \in K),$$

en tanto que (59) y (60) implicarán que

$$h(t) < f(t) + \varepsilon \qquad (t \in K). \tag{62}$$

Finalmente, (58) se deduce a partir de (61) y de (62).

El Teorema 7.32 no se cumple para álgebras complejas. En el Ejercicio 21 se da un ejemplo de este hecho. Sin embargo, se cumple la conclusión del teorema, aun para álgebras complejas, si se impone una condición suplementaria a $\mathscr{A}$, esto es, que sea *auto-adjunta*. Esto significa que para toda $f \in \mathscr{A}$ su conjugada compleja $\bar{f}$ debe pertenecer también a $\mathscr{A}$; $\bar{f}$ está definida por $\bar{f}(x) = \overline{f(x)}$.

7.33 Teorema *Supongamos que $\mathscr{A}$ es un álgebra auto-adjunta de funciones complejas continuas en un conjunto compacto K, que $\mathscr{A}$ separa puntos en K y no desaparece en ningún punto de K. La cerradura uniforme $\mathscr{B}$ de $\mathscr{A}$ consta de todas las funciones complejas continuas en K. En otras palabras, $\mathscr{A}$ es denso en $\mathscr{C}(K)$.*

Demostración Sea $\mathscr{A}_R$ el conjunto de todas las funciones reales en K que pertenecen a $\mathscr{A}$.

Si $f \in \mathscr{A}$ y $f = u + iv$, con u, v reales, será $2u = f + \bar{f}$, y como $\mathscr{A}$ es auto-adjunta, vemos que $u \in \mathscr{A}_R$. Si $x_1 \neq x_2$, existe $f \in \mathscr{A}$, tal que $f(x_1) = 1$, $f(x_2) = 0$; por tanto, $0 = u(x_2) \neq u(x_1) = 1$, lo que demuestra que $\mathscr{A}_R$ separa puntos en K. Si $x \in K$, entonces $g(x) \neq 0$ para algún $g \in \mathscr{A}$, y hay un número complejo λ, tal que $\lambda g(x) > 0$; si $f = \lambda g, f = u + iv$, se deduce que $u(x) > 0$; por tanto, $\mathscr{A}_R$ no desaparece en ningún punto de K.

Así, pues, $\mathscr{A}_R$ satisface las hipótesis del Teorema 7.32. Se deduce que toda función real continua en K está en la cerradura uniforme de $\mathscr{A}_R$, por lo que está en $\mathscr{B}$. Si f es una función compleja continua en $K, f = u + iv$, entonces $u \in \mathscr{B}$, $v \in \mathscr{B}$, y, por tanto, $f \in \mathscr{B}$, lo que completa la demostración.

EJERCICIOS

1. Demostrar que toda sucesión uniformemente convergente de funciones acotadas, es uniformemente acotada.
2. Si $\{f_n\}$ y $\{g_n\}$ convergen uniformemente en E, demostrar que $\{f_n + g_n\}$ convergen uniformemente en E. Si, además $\{f_n\}$ y $\{g_n\}$ son sucesiones de funciones acotadas, demostrar que $\{f_n g_n\}$ convergen uniformemente en E.
3. $\{f_n g_n\}$ no converja uniformemente en E (desde luego, $\{f_n g_n\}$ deben converger en E).
4. Considerar

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{1 + n^2 x}.$$

¿Para qué valores de x converge la serie absolutamente? ¿En qué intervalos deja de converger uniformemente? ¿Es f continua dondequiera que converja la serie? ¿Es acotada f?

5. Sea

$$f_n(x) = \begin{cases} 0 & \left(x < \dfrac{1}{n+1}\right), \\ \operatorname{sen}^2 \dfrac{\pi}{x} & \left(\dfrac{1}{n+1} \leq x \leq \dfrac{1}{n}\right), \\ 0 & \left(\dfrac{1}{n} < x\right). \end{cases}$$

Demostrar que $\{f_n\}$ converge hacia una función continua, pero no uniformemente. Utilizar la serie Σf_n para demostrar que la convergencia absoluta, aun para todo x, no implica convergencia uniforme.

6. Demostrar que la serie

$$\sum_{n=1}^{\infty} (-1)^n \frac{x^2+n}{n^2}$$

converge uniformemente en todo intervalo acotado, pero no converge absolutamente para cualquier valor de x.

7. Para $n = 1, 2, 3, \ldots$, y x real, hacer

$$f_n(x) = \frac{x}{1+nx^2}.$$

Demostrar que $\{f_n\}$ converge uniformemente hacia una función f, y que la ecuación

$$f'(x) = \lim_{n \to \infty} f_n'(x)$$

es correcta si $x \neq 0$, pero falsa si $x = 0$.

8. Si

$$I(x) = \begin{cases} 0 & (x \leq 0), \\ 1 & (x > 0), \end{cases}$$

si $\{x_n\}$ es una sucesión de puntos distintos de (a,b), y si $\Sigma|c_n|$ converge, demostrar que la serie

$$f(x) = \sum_{n=1}^{\infty} c_n I(x - x_n) \qquad (a \leq x \leq b)$$

converge uniformemente, que f es continua para todo $x \neq x_n$.

9. Sea $\{f_n\}$ una sucesión de funciones continuas que converge uniformemente hacia una función f en un conjunto E. Demostrar que

$$\lim_{n \to \infty} f_n(x_n) = f(x)$$

para toda sucesión de puntos $x_n \in E$ tales que $x_n \to x$, y $x \in E$. ¿Es cierto el recíproco?

10. Si (x) representa la parte fraccionaria de un número real x (ver, para la definición, el Ejercicio 16, Cap. 4), considerar la función

$$f(x) = \sum_{n=1}^{\infty} \frac{(nx)}{n^2} \qquad (x \text{ real}).$$

Hallar todas las discontinuidades de f, y demostrar que forman un conjunto denso numerable. Demostrar que f es, sin embargo, integrable según Riemann en todo intervalo acotado.

11. Suponer que $\{f_n\}$, $\{g_n\}$ están definidas en E, y

(*a*) Σf_n tiene sumas parciales uniformemente acotadas.

(*b*) $g_n \to 0$ uniformemente en E.

(*c*) $g_1(x) \geq g_2(x) \geq g_3(x) \geq \cdots$ para todo $x \in E$. Entonces, $\Sigma f_n g_n$ converge uniformemente en E. *Sugerencia*: Comparar con el Teorema 3.42.

12. Supóngase que g y $f_n (n = 1, 2, 3, \ldots)$ están definidas sobre $(0, \infty)$, son integrables según Riemann sobre $[t, T]$ siempre que $0 < t < T < \infty$, $|f_n| \leq g, f_n \to f$ uniformemente sobre cada subconjunto compacto de $[0, \infty]$, y que

$$\int_0^{\infty} g(x)\, dx < \infty.$$

Demostrar que

$$\lim_{n\to\infty} \int_0^{\infty} f_n(x)\, dx = \int_0^{\infty} f(x)\, dx.$$

(Para las definiciones que se presentan, véanse los Ejercicios 7 y 8 del Cap. 6.)

Esta es una forma bastante débil del teorema de Lebesgue de la convergencia dominada (Teorema 11.32). En el contexto de la integral de Riemann, puede incluso reemplazarse la convergencia uniforme, por la convergencia puntual, si se admite que $f \in \mathscr{R}$. (Véanse los artículos de F. Cunningham en el *Math. Mag.*, tomo 40, 1967, páginas 179 a 186, y H. Kestelman en la *Amer. Math. Monthly*, tomo 77, 1970, páginas 182 a la 187.)

13. Supóngase que $\{f_n\}$ es una sucesión de funciones monótonas crecientes sobre R^1, con $0 \leq f_n(x) \leq 1$ para todo x y todo n.

(*a*) Demostrar que hay una función f y una sucesión $\{n_k\}$ tales que

$$f(x) = \lim_{k\to\infty} f_{n_k}(x)$$

para cada $x \in R^1$. (La existencia de tal sucesión puntualmente convergente se conoce comúnmente como *teorema de selección de Helly*.)

(*b*) Si además, f es continua, demostrar que $f_{n_k} \to f$ uniformemente sobre R^1.

Sugerencia: (i) Alguna subsucesión $\{f_{n_i}\}$ converge en todos los puntos racionales r, por ejemplo, hacia $f(r)$. (ii) Definir $f(x)$, para cualquier $x \in R^1$, como el sup $f(r)$, en donde el sup se toma sobre todos los $r \leq x$. (iii) Mostrar que $f_{n_i}(x) \to f(x)$ en cada x en el cual f es continua. (Es decir, donde la monotonía se usa con más rigor) (iv) Una subsucesión de $\{f_{n_i}\}$ converge en cada punto de discontinuidad de f por que a lo más hay un número numerable de tales puntos. Esto demuestra (a). Para demostrar (b) conviene modificar la demostración de (iii) apropiadamente.

14. Sea f una función real continua sobre R^1, que tiene las siguientes propiedades: $0 \leq f(t) \leq 1$, $f(t+2) = f(t)$ para cada t, y

$$f(t)=\begin{cases}0 & (0\leq t\leq \frac{1}{3})\\ 1 & (\frac{2}{3}\leq t\leq 1).\end{cases}$$

Si $\Phi(t) = (x(t), y(t))$, donde

$$x(t)=\sum_{n=1}^{\infty}2^{-n}f(3^{2n-1}t), \qquad y(t)=\sum_{n=1}^{\infty}2^{-n}f(3^{2n}t).$$

Demostrar que Φ es *continua* y que Φ mapea $I = [0, 1]$ *sobre* el cuadrado unitario $I^2 \subset R^2$. En realidad se tiene que mostrar que Φ mapea el conjunto de Cantor sobre I^2.

Sugerencia: Cada $(x_0, y_0) \in I^2$ tiene la forma

$$x_0=\sum_{n=1}^{\infty}2^{-n}a_{2n-1}, \qquad y_0=\sum_{n=1}^{\infty}2^{-n}a_{2n}$$

en donde cada a_i es 0 ó 1. Si

$$t_0=\sum_{i=1}^{\infty}3^{-i-1}(2a_i)$$

mostrar que $f(3^k t_0) = a_k$, y por consiguiente, que $x(t_0) = x_0$, $y(t_0) = y_0$.

(Este sencillo ejemplo de la llamada "curva que llena el espacio" se debe a I. J. Schoenberg, *Bull. A.M.S.*, tomo 44, 1938, página 519.)

15. Supóngase que f es una función real continua sobre R^1, $f_n(t) = f(nt)$ para $n = 1, 2, 3, \ldots$, y $\{f_n\}$ es equicontinua sobre $[0, 1]$. ¿Qué puede decirse de f?

16. Supóngase que $\{f_n\}$ es una sucesión equicontinua de funciones sobre un conjunto compacto K, y que $\{f_n\}$ converge puntualmente sobre K. Demostrar que $\{f_n\}$ converge uniformemente sobre K.

17. Las nociones de convergencia uniforme y equicontinuidad se definen para mapeos en cualquier espacio métrico. Demostrar que los Teoremas 7.9 y 7.12 son válidos para mapeos en cualquier espacio métrico, y que los Teoremas 7.8 y 7.11 son válidos para mapeos en cualquier espacio métrico completo. Y por último, demostrar que para funciones vectoriales, es decir, aplicaciones en cualquier R^k, los Teoremas 7.10, 7.16, 7.17, 7.24 y 7.25 son válidos.

18. Sea $\{f_n\}$ una sucesión uniformemente acotada de funciones que son integrables según Riemann en $[a,b]$ y hagamos

$$F_n(x) = \int_a^x f_n(t)\,dt \qquad (a \le x \le b).$$

Demostrar que existe una subsucesión $\{F_{n_k}\}$ que converge uniformemente en $[a\ b]$.

19. Sea K un espacio métrico compacto y S un subconjunto de $\mathscr{C}(K)$. Demostrar que S es compacto (con respecto a la métrica que se definió en la Sec. 7.14) si, y solo si S es uniformemente cerrado, acotado puntualmente y equicontinuo. (Si S no es equicontinuo, entonces contiene una sucesión que no tiene una subsucesión equicontinua, por consiguiente no tiene subsucesión que converja uniformemente sobre K.)

20. Si f es continua en $[0, 1]$ y si

$$\int_0^1 f(x)x^n\,dx = 0 \qquad (n = 0, 1, 2, \ldots),$$

entonces, $f(x) = 0$ en $[0, 1]$. *Sugerencia*: La integral del producto de f por cualquier polinomio es cero. Utilizar el teorema de Weierstrass para demostrar que $\int_0^1 f^2(x)\,dx = 0$.

21. Sea K el círculo unitario en el plano complejo (esto es, el conjunto de todos los z con $|z| = 1$) y sea $\mathscr{A}$ el álgebra de todas las funciones de la forma

$$f(e^{i\theta}) = \sum_{n=0}^{N} c_n e^{in\theta} \qquad (\theta \text{ real}).$$

Entonces, $\mathscr{A}$ separa puntos en K y no desaparece en ningún punto de K, pero sin embargo hay funciones continuas en K que no están en la cerradura uniforme de $\mathscr{A}$. *Sugerencia*: Para cada $f \in \mathscr{A}$.

$$\int_0^{2\pi} f(e^{i\theta})e^{i\theta}\,d\theta = 0,$$

y esto es también cierto para cada f en la cerradura de $\mathscr{A}$.

22. Sea $f \in \mathscr{R}(\alpha)$ sobre $[a,b]$. Demostrar que hay polinomios P_n tales que

$$\lim_{n\to\infty} \int_a^b |f - P_n|^2\,d\alpha = 0.$$

(Comparar esto con el Ejercicio 12 del Cap. 6.)

23. Hágase $P_0 = 0$, y para $n = 0, 1, 2, \ldots$, defínase

$$P_{n+1}(x) = P_n(x) + \frac{x^2 - P_n^2(x)}{2}.$$

Demostrar que

$$\lim_{n\to\infty} P_n(x) = |x|,$$

uniformemente en $[-1, 1]$.

(Esto hace posible demostrar el teorema de Stone-Weierstrass sin probar de antemano el Teorema 7.26).

Sugerencia: Usar la identidad

$$|x| - P_{n+1}(x) = [|x| - P_n(x)]\left[1 - \frac{|x| + P_n(x)}{2}\right]$$

para demostrar que $0 \leq P_n(x) \leq P_{n+1}(x) \leq |x|$ si $|x| \leq 1$, y que

$$|x| - P_n(x) \leq |x|\left(1 - \frac{|x|}{2}\right)^n < \frac{2}{n+1}$$

si $|x| \leq 1$.

24. Sea X un espacio métrico, con métrica d. Si se fija un punto $a \in X$ y se asigna a cada $p \in X$ la función f_p definida por

$$f_p(x) = d(x, p) - d(x, a) \qquad (x \in X).$$

Demostrar que $|f_p(x)| \leq d(a,p)$ para todo $x \in X$, y que por lo tanto, $f_p \in \mathscr{C}(X)$.

Demostrar después que

$$\|f_p - f_q\| = d(p, q)$$

para todo $p,q \in X$.

Si $\Phi(p) = f_p$ se deduce que Φ es una *isometría* (una aplicación o mapeo que preserva distancias) de X sobre $\Phi(X) \subset \mathscr{C}(X)$.

Sea Y la cerradura de $\Phi(X)$ en $\mathscr{C}(X)$. Mostrar que Y es completo.

Conclusión: *X es isométrico a un subconjunto denso de un espacio métrico completo Y.*

(Para una demostración diferente de esto, véase el Ejercicio 24 del Cap. 3.)

25. Supongamos que ϕ es una función real continua acotada en la banda definida por $0 \leq x \leq 1$; $-\infty < y < \infty$. Demostrar que el problema con valores iniciales

$$y' = \phi(x, y), \qquad y(0) = c$$

tiene una solución. (Obsérvese que las hipótesis de este teorema de existencia son menos restrictivas que las del correspondiente teorema de unicidad; ver el Ejercicio 2, Cap. 5).

Sugerencia: Fijar n. Para $i = 0,\ldots, n$, hacer $x_i = i/n$. Sea f_n una función continua en $[0, 1]$ tal que $f_n(0) = c$,

$$f_n'(t) = \phi(x_i, f_n(x_i)) \qquad \text{si } x_i < t < x_{i+1},$$

y hagamos

$$\Delta_n(t) = f_n'(t) - \phi(t, f_n(t)),$$

excepto en los puntos x_i, donde $\Delta_n(t) = 0$. Será

$$f_n(x) = c + \int_0^x [\phi(t, f_n(t)) + \Delta_n(t)]\, dt.$$

Selecciónese $M < \infty$ de tal forma que $|\phi| \leq M$. Verificar las afirmaciones que siguen:

(a) $|f_n'| \leq M$, $|\Delta_n| \leq 2M$, $\Delta_n \in \mathscr{R}$, y $|f_n| \leq |c| + M = M_1$, digamos, sobre $[0, 1]$, para todo n.
(b) $\{f_n\}$ es equicontinua sobre $[0,1]$, debido a que $|f_n'| \leq M$.
(c) Alguna $\{f_{n_k}\}$ converge hacia alguna f, uniformemente sobre $[0, 1]$.
(d) Como ϕ es uniformemente continua sobre el rectángulo $0 \leq x \leq 1$, $|y| \leq M_1$,

$$\phi(t, f_{n_k}(t)) \to \phi(t, f(t))$$

uniformemente sobre $[0, 1]$.
(e) $\Delta_n(t) \to 0$ uniformemente sobre $[0, 1]$, debido a que

$$\Delta_n(t) = \phi(x_i, f_n(x_i)) - \phi(t, f_n(t))$$

en (x_i, x_{i+1}).
(f) Por consiguiente

$$f(x) = c + \int_0^x \phi(t, f(t))\, dt.$$

Esta f es una solución del problema planteado.

26. Demostrar un teorema de existencia análogo para el problema con valores iniciales

$$\mathbf{y}' = \mathbf{\Phi}(x, \mathbf{y}), \qquad \mathbf{y}(0) = \mathbf{c},$$

en el cual $\mathbf{c} \in R^k$, $\mathbf{y} \in R^k$, y $\mathbf{\Phi}$ es un mapeo continuo acotado, de la parte de R^{k+1} definida por $0 \leq x \leq 1$, $\mathbf{y} \in R^k$ en R^k. (Compararlo con el Ejercicio 28 del Cap. 5.) *Sugerencia*: Usar la versión vectorial del Teorema 7.25.

8

ALGUNAS FUNCIONES ESPECIALES

SERIES DE POTENCIAS

En esta sección deduciremos algunas propiedades de las funciones que están representadas por series de potencias, es decir, funciones de la forma

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \tag{1}$$

o, más generalmente,

$$f(x) = \sum_{n=0}^{\infty} c_n (x - a)^n. \tag{2}$$

Se les llama *funciones analíticas.*

Nos limitaremos a valores reales de x. En lugar de círculos de convergencia (ver el Teorema 3.39) encontraremos, por tanto, intervalos de convergencia.

Si (1) converge para todo x en $(-R,R)$, para algún $R > 0$ (R puede ser $+\infty$), decimos que f está desarrollada en serie de potencias en torno del punto $x = 0$. De igual modo; si (2) converge para $|x - a| < R$, se dice que f está desarrollada en una serie de potencias en torno del punto $x = a$. Por conveniencia tomaremos, a menudo $a = 0$ sin pérdida de generalidad.

8.1 Teorema *Supongamos que la serie*

(3) $$\sum_{n=0}^{\infty} c_n x^n$$

converge para $|x| < R$, *y definamos*

(4) $$f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad (|x| < R).$$

Entonces, (3) *converge uniformemente en* $[-R + \varepsilon, R - \varepsilon]$, *para cualquier* ε *prefijado. La función* f *es continua y diferenciable en* $(-R,R)$, *y*

(5) $$f'(x) = \sum_{n=1}^{\infty} nc_n x^{n-1} \qquad (|x| < R).$$

Demostración Sea $\varepsilon > 0$ dado. Para $|x| \leq R - \varepsilon$, tenemos

$$|c_n x^n| \leq |c_n(R-\varepsilon)^n|;$$

y como

$$\Sigma c_n(R-\varepsilon)^n$$

converge absolutamente (toda serie de potencias converge absolutamente en el interior de su intervalo de convergencia, por el criterio de la raíz), el Teorema 7.10 demuestra la convergencia uniforme de (3) en $[-R + \varepsilon, R - \varepsilon]$.

Como $\sqrt[n]{n} \to 1$ cuando $n \to \infty$, tenemos

$$\limsup_{n\to\infty} \sqrt[n]{n|c_n|} = \limsup_{n\to\infty} \sqrt[n]{|c_n|},$$

de modo que las series (4) y (5) tienen el mismo intervalo de convergencia.

Como (5) es una serie de potencias, converge uniformemente en $[-R + \varepsilon, R - \varepsilon]$ para todo $\varepsilon > 0$, y podemos aplicar el Teorema 7.17 (para series en lugar de sucesiones). Se deduce que (5) se cumple si $|x| \leq R - \varepsilon$.

Pero, dado cualquier x tal que $|x| < R$, podemos hallar un $\varepsilon > 0$ tal que $|x| < R - \varepsilon$, lo que demuestra que se cumple (5) para $|x| < R$.

De la existencia de f' se deduce la continuidad de f (Teorema 5.2).

Corolario *Con la hipótesis del Teorema* 8.1, f *tiene derivadas de todos los órdenes en* $(-R,R)$, *que vienen dadas por*

(6) $$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)c_n x^{n-k}.$$

En particular,

(7) $$f^{(k)}(0) = k!c_k \qquad (k = 0, 1, 2, \ldots).$$

(Aquí $f^{(0)}$ significa f, y $f^{(k)}$ es la derivada de orden k de f, para $k = 1, 2, 3, \ldots$)

Demostración La ecuación (6) se deduce aplicando el Teorema 8.1 sucesivamente a $f, f', f'', \ldots$ Haciendo $x = 0$ en (6), obtenemos (7).

La fórmula (7) es muy interesante. Demuestra, por un lado, que los coeficientes del desarrollo en serie de potencias de f están determinados por los valores de f y de sus derivadas en un punto. Por otro lado, si están dados los coeficientes, los valores de las derivadas de f en el centro del intervalo de convergencia se ven inmediatamente en la serie de potencias.

Obsérvese, sin embargo, que aunque una función f tenga derivadas de todos los órdenes, la serie $\Sigma c_n x^n$, donde c_n está calculado según (7), no converge necesariamente hacia $f(x)$ para cualquier $x \neq 0$. En este caso, no puede desarrollarse f en una serie de potencias en torno de $x = 0$. Porque tenemos que $f(x) = \Sigma a_n x^n$, debería ser

$$n!a_n = f^{(n)}(0);$$

y, por tanto $a_n = c_n$. Un ejemplo de este caso, se dará en el Ejercicio 1.

Si la serie (3) converge en un extremo, por ejemplo $x = R$, f es continua no solo en $(-R,R)$, sino también en $x = R$, lo que se deduce del teorema de Abel (para sencillez de notación tomamos $R = 1$).

8.2 Teorema *Supongamos que* Σc_n *converge. Hagamos*

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad (-1 < x < 1).$$

Será

(8) $$\lim_{x\to 1} f(x) = \sum_{n=0}^{\infty} c_n.$$

Demostración Sea $s_n = c_0 + \cdots + c_n$; $s_{-1} = 0$. Entonces

$$\sum_{n=0}^{m} c_n x^n = \sum_{n=0}^{m} (s_n - s_{n-1})x^n = (1-x)\sum_{n=0}^{m-1} s_n x^n + s_m x^m.$$

Para $|x| < 1$, hagamos $m \to \infty$ y obtendremos

(9)
$$f(x) = (1-x)\sum_{n=0}^{\infty} s_n x^n.$$

Supongamos $s = \lim_{n\to 0} s_n$. Sea $\varepsilon > 0$ dado. Elijamos N tal que $n > N$ implique

$$|s - s_n| < \frac{\varepsilon}{2}.$$

Como

$$(1-x)\sum_{n=0}^{\infty} x^n = 1 \qquad (|x| < 1),$$

de (9) obtendremos

$$|f(x) - s| = \left|(1-x)\sum_{n=0}^{\infty}(s_n - s)x^n\right| \le (1-x)\sum_{n=0}^{N}|s_n - s|\,|x|^n + \frac{\varepsilon}{2} \le \varepsilon$$

si $x > 1 - \delta$, para algún $\delta > 0$ convenientemente elegido, lo que implica (8).

Como aplicación, demostremos el Teorema 3.51, que dice: *si* Σa_n, Σb_n, Σc_n, *convergen hacia* A, B, C, *y si* $c_n = a_0 b_n + \cdots + a_n b_0$, *será* $C = AB$. Sea

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \qquad g(x) = \sum_{n=0}^{\infty} b_n x^n, \qquad h(x) = \sum_{n=0}^{\infty} c_n x^n,$$

para $0 \le x \le 1$. Para $x < 1$, estas series convergen absolutamente y, por tanto, pueden nultiplicarse de acuerdo con la Definición 3.48; una vez realizada la multiplicación vemos que

(10)
$$f(x)\cdot g(x) = h(x) \qquad (0 \le x < 1).$$

Por el Teorema 8.2,

(11)
$$f(x) \to A, \qquad g(x) \to B, \qquad h(x) \to C$$

cuando $x \to 1$. Las ecuaciones (10) y (11) implican que $AB = C$.

Ahora necesitamos un teorema sobre la inversión del orden de la suma. (Véanse Ej. 2 y 3.)

8.3 Teorema *Dada una doble sucesión* $\{a_{ij}\}$, $i = 1, 2, 3, \ldots, j = 1, 2, 3, \ldots$, *supongamos que*

$$\sum_{j=1}^{\infty} |a_{ij}| = b_i \qquad (i = 1, 2, 3, \ldots) \tag{12}$$

y que Σb_i *converge. Entonces,*

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}. \tag{13}$$

Demostración Podríamos establecer (13) por un procedimiento directo similar (aunque más complicado) al utilizado en el Teorema 3.55. Sin embargo, el método siguiente parece más interesante.

Sea E un conjunto numerable, constituido por los puntos $x_0, x_1, x_2, \ldots$, y supongamos que $x_n \to x_0$ cuando $n \to \infty$. Definamos

$$f_i(x_0) = \sum_{j=1}^{\infty} a_{ij} \qquad (i = 1, 2, 3, \ldots), \tag{14}$$

$$f_i(x_n) = \sum_{j=1}^{n} a_{ij} \qquad (i, n = 1, 2, 3, \ldots), \tag{15}$$

$$g(x) = \sum_{i=1}^{\infty} f_i(x) \qquad (x \in E). \tag{16}$$

Ahora (14) y (15) juntamente con (12), demuestran que toda f_i es continua en x_0. Como $|f_i(x)| \le b_i$ para $x \in E$ (16) converge uniformemente, de modo que g es continua en x_0 (Teorema 7.11). Se deduce que

$$\begin{aligned}
\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} &= \sum_{i=1}^{\infty} f_i(x_0) = g(x_0) = \lim_{n\to\infty} g(x_n) \\
&= \lim_{n\to\infty} \sum_{i=1}^{\infty} f_i(x_n) = \lim_{n\to\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{n} a_{ij} \\
&= \lim_{n\to\infty} \sum_{j=1}^{n} \sum_{i=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}.
\end{aligned}$$

8.4 Teorema *Supongamos que*

$$f(x) = \sum_{n=0}^{\infty} c_n x^n,$$

convergiendo la serie en $|x| < R$. Si $-R < a < R$, *f puede ser desarrollada en serie de potencias alrededor del punto* $x = a$, *que converge en* $|x - a| < R - |a|$, *y*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \qquad (|x-a| < R - |a|). \tag{17}$$

Esto es una generalización del Teorema 5.15 y se conoce también como teorema de Taylor.

Demostración Tenemos

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} c_n[(x-a)+a]^n \\ &= \sum_{n=0}^{\infty} c_n \sum_{m=0}^{n} \binom{n}{m} a^{n-m}(x-a)^m \\ &= \sum_{m=0}^{\infty} \left[\sum_{n=m}^{\infty} \binom{n}{m} c_n a^{n-m}\right](x-a)^m. \end{aligned}$$

que es el desarrollo alrededor del punto $x = a$ buscado. Para demostrar su validez, tenemos que justificar el cambio realizado en el orden de la suma. El Teorema 8.3 demuestra que es permisible si

$$\sum_{n=0}^{\infty} \sum_{m=0}^{n} \left| c_n \binom{n}{m} a^{n-m}(x-a)^m \right| \tag{18}$$

converge. Pero (18) es lo mismo que

$$\sum_{n=0}^{\infty} |c_n| \cdot (|x-a| + |a|)^n, \tag{19}$$

y (19) converge si $|x - a| + |a| < R$.

Finalmente, la forma de los coeficientes de (17) se deduce de (7).

Se habrá observado que (17) puede converger realmente en un intervalo mayor que el dado por $|x - a| < R - |a|$.

Si dos series de potencias convergen hacia la misma función en $(-R,R)$ (7) demuestra que las dos series deben ser idénticas, esto es, deben tener los mismos coeficientes. Es interesante observar que puede llegarse a la misma conclusión a partir de hipótesis mucho menos restrictivas.

8.5 Teorema *Supongamos que las series* $\Sigma a_n x^n$ *y* $\Sigma b_n x^n$ *convergen en el*

segmento $S = (-R,R)$. Sea E el conjunto de todos los $x \in S$, tales que

(20)
$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n.$$

Si E tiene un punto límite en S, entonces $a_n = b_n$ para $n = 0, 1, 2, \ldots$ Por tanto, se cumple (20) *para todo $x \in S$.*

Demostración Hagamos $c_n = a_n - b_n$ y

(21)
$$f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad (x \in S).$$

Entonces $f(x) = 0$ sobre E.

Sea A el conjunto de todos los puntos límite de E en S, y supóngase que B consta de todos los demás puntos de S. Se ve claramente de la definición de "punto límite" que B es abierto. Supóngase también que se puede demostrar que A es abierto. Entonces A y B son conjuntos abiertos ajenos. Por esto, son separados (Definición 2.45). Como $S = A \cup B$, y S es conexo, uno de los A y B debe ser vacío. Por hipótesis, A no es vacío. En consecuencia B es vacío, y $A = S$. Debido a que f es continua en S, $A \subset E$. Entonces $E = S$, y (7) muestra que $c_n = 0$ para $n = 0, 1, 2, \ldots$ que es la conclusión buscada.

Hemos demostrado, así, que A es abierto. Si $x_0 \in A$, el Teorema 8.4 demuestra que

(22)
$$f(x) = \sum_{n=0}^{\infty} d_n(x - x_0)^n \qquad (|x - x_0| < R - |x_0|).$$

Deberá ser $d_n = 0$ para todo n. De otro modo, sea k el menor entero no negativo, tal que $d_k \neq 0$. Entonces

(23)
$$f(x) = (x - x_0)^k g(x) \qquad (|x - x_0| < R - |x_0|),$$

donde

(24)
$$g(x) = \sum_{m=0}^{\infty} d_{k+m}(x - x_0)^m.$$

Como g es continua en x_0 y

$$g(x_0) = d_k \neq 0,$$

existe un $\delta > 0$, tal que $g(x) \neq 0$ si $|x - x_0| < \delta$. Se deduce de (23)

que $f(x) \neq 0$ si $0 < |x - x_0| < \delta$. Pero esto contradice el hecho de ser x_0 un punto límite de E.

Así, pues, $d_n = 0$ para todo n, de modo que $f(x) = 0$ para todo x para el que se cumpla (22), esto es, en una vecindad de x_0, lo que prueba que A es abierto, y completa la demostración.

LAS FUNCIONES EXPONENCIAL Y LOGARÍTMICA

Definiremos

(25) $$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

El criterio de la razón demuestra que esta serie converge para todo número complejo z. Aplicando el Teorema 3.50 a la multiplicación de series absolutamente convergentes, obtenemos

$$\begin{aligned} E(z)E(w) &= \sum_{n=0}^{\infty} \frac{z^n}{n!} \sum_{m=0}^{\infty} \frac{w^m}{m!} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{z^k w^{n-k}}{k!(n-k)!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} = \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!}, \end{aligned}$$

lo que nos da la importante fórmula de la adición

(26) $$E(z + w) = E(z)E(w) \qquad (z,\ w \text{ complejas})$$

Una consecuencia es que,

(27) $$E(z)E(-z) = E(z - z) = E(0) = 1 \qquad (z \text{ compleja})$$

lo que demuestra que $E(z) \neq 0$ para todo z. Por (25), $E(x) > 0$ si $x > 0$; por tanto (27) muestra que $E(x) > 0$ para todo x real. Por (25), $E(x) \to +\infty$ cuando $x \to +\infty$; por lo que (27) muestra que $E(x) \to 0$ cuando $x \to -\infty$ a lo largo del eje real. Por (25), $0 < x < y$ implica que $E(x) < E(y)$; por (27) se deduce que $E(-y) < E(-x)$, por lo cual E es estrictamente creciente en todo el eje real.

La fórmula de la adición demuestra también que

(28) $$\lim_{h=0} \frac{E(z+h) - E(z)}{h} = E(z) \lim_{h=0} \frac{E(h) - 1}{h} = E(z);$$

la última igualdad se deduce directamente de (25).

La iteración de (26) da

$$E(z_1 + \cdots + z_n) = E(z_1) \cdots E(z_n). \tag{29}$$

Tomemos $z_1 = \cdots = z_n = 1$. Como $E(1) = e$, siendo e el número definido en la Definición 3.30, obtenemos

$$E(n) = e^n \qquad (n = 1, 2, 3, \ldots). \tag{30}$$

Si $p = n/m$, siendo n, m enteros positivos, será

$$[E(p)]^m = E(mp) = E(n) = e^n, \tag{31}$$

de modo que

$$E(p) = e^p \qquad (p > 0, p \text{ racional}). \tag{32}$$

De (27) se deduce que $E(-p) = e^{-p}$ si p es positivo y racional. Así, pues, se cumple (32) para todo número racional p.

En el Ejercicio 6, capítulo 1 planteamos la definición

$$x^y = \sup x^p, \tag{33}$$

donde el sup se toma sobre todo racional p, tal que $p < y$, para cualquier número real y, y $x > 1$. Si definimos, pues, para todo x real

$$e^x = \sup e^p \qquad (p < x, p \text{ racional}) \tag{34}$$

las propiedades de monotonía y continuidad de E, junto con (32), demuestran que

$$E(x) = e^x \tag{35}$$

para todo x real. La ecuación (35) explica por qué a E se le llama función exponencial.

La notación exp (x). se usa más que la e^x, especialmente cuando x es una expresión complicada.

Se puede utilizar perfectamente (35) en lugar de (34) como definición de e^x; (35) es un punto de partida mucho más conveniente para el estudio de las propiedades de e^x. Veremos ahora que también puede sustituirse (33) por una definición más conveniente [ver (43)].

Pasando, ahora, a la notación habitual, e^x, en lugar de $E(x)$, resumiremos lo que hemos demostrado hasta ahora.

8.6 Teorema *Supongamos* e^x *definida en* R^1 *por* (35) *y* (25). *Entonces*

(*a*) e^x *es continua y diferenciable para todo* x;

(*b*) $(e^x)' = e^x$;

(*c*) e^x *es función estrictamente creciente de* x, *y* $e^x > 0$;

(*d*) $e^{x+y} = e^x e^y$;

(*e*) $e^x \to +\infty$ *cuando* $x \to +\infty$, $e^x \to 0$ *cuando* $x \to -\infty$;

(*f*) $\lim_{x\to+\infty} x^n e^{-x} = 0$, *para todo* n.

Hemos demostrado ya de (*a*) a (*e*); (25) demuestra que

$$e^x > \frac{x^{n+1}}{(n+1)!}$$

para $x > 0$, así, que

$$x^n e^{-x} < \frac{(n+1)!}{x},$$

y se deduce (*f*). El apartado (*f*) demuestra que e^x tiende hacia $+\infty$ «más de prisa» que cualquier potencia de x, cuando $x \to +\infty$.

Como E es estrictamente creciente y diferenciable en F^1, tiene una función inversa L que también es estrictamente creciente y diferenciable y cuyo dominio de definición es $E(R^1)$, esto es, el conjunto de todos los números positivos. L está definido por

(36) $$E(L(y)) = y \qquad (y > 0),$$

o, de forma equivalente, por

(37) $$L(E(x)) = x \qquad (x \text{ real}).$$

Diferenciando (37), obtenemos (compárese con el Teorema 5.5)

$$L'(E(x)) \cdot E(x) = 1.$$

Haciendo $y = E(x)$, nos da

(38) $$L'(y) = \frac{1}{y} \qquad (y > 0).$$

Tomando $x = 0$ en (37), vemos que $L(1) = 0$. Por tanto, (38) implica

(39) $$L(y) = \int_1^y \frac{dx}{x}.$$

Muy frecuentemente, se toma (39) como punto de partida de la teoría de los logaritmos y de la función exponencial. Haciendo $u = E(x)$, $v = E(y)$, (26) da

$$L(uv) = L(E(x) \cdot E(y)) = L(E(x + y)) = x + y,$$

de modo que

$$L(uv) = L(u) + L(v) \qquad (u > 0,\ v > 0). \tag{40}$$

Esto demuestra que L tiene la propiedad conocida que hace de los logaritmos un instrumento de cálculo. La notación acostumbrada para $L(x)$ es, naturalmente, $\log x$.

En cuanto al comportamiento de $\log x$ cuando $x \to +\infty$ y cuando $x \to 0$, el Teorema 8.6(e) demuestra que

$$\begin{aligned} \log x &\to +\infty \qquad \text{si } x \to +\infty, \\ \log x &\to -\infty \qquad \text{si } x \to 0. \end{aligned}$$

Se ve fácilmente que

$$x^n = E(nL(x)) \tag{41}$$

si $x > 0$ y n es un entero. De igual modo, si m es un entero positivo, tenemos

$$x^{1/m} = E\left(\frac{1}{m} L(x)\right), \tag{42}$$

pues cada término de (42), elevado a la m^a potencia, se convierte en el correspondiente en (37). Combinando (41) y (42) obtenemos

$$x^\alpha = E(\alpha L(x)) = e^{\alpha \log x} \tag{43}$$

para todo número racional α.

Definiremos, ahora, x^α para todo α real y todo $x > 0$, por (43). La continuidad y monotonía de E y L demuestra que esta definición conduce al mismo resultado que la dada anteriormente. Los resultados establecidos en el Ejercicio 6 del capítulo 1 son consecuencias inmediatas de (43).

Si diferenciamos (43), por el Teorema 5.5, obtenemos

$$(x^\alpha)' = E(\alpha L(x)) \cdot \frac{\alpha}{x} = \alpha x^{\alpha - 1}. \tag{44}$$

Obsérvese que hemos utilizado previamente (44) solo para valores enteros de α, en cuyo caso se deduce (44) fácilmente del Teorema 5.3(b). El demostrarlo directamente a partir de la definición de derivada, si x^α está definida por (33) y α es irracional, es muy engorroso.

La conocida fórmula de integración para x^α se deduce de (44) si $\alpha \neq -1$, y de (38) si $\alpha = -1$. Demostraremos una nueva propiedad de $\log x$, a saber:

$$\lim_{x\to+\infty} x^{-\alpha}\log x = 0 \tag{45}$$

para todo $\alpha > 0$. Esto es, $\log x \to +\infty$ «más lentamente» que cualquier potencia positiva de x, cuando $x \to +\infty$.

Porque si $0 < \varepsilon < \alpha$, y $x > 1$, será

$$\begin{aligned} x^{-\alpha}\log x &= x^{-\alpha}\int_1^x t^{-1}\,dt < x^{-\alpha}\int_1^x t^{\varepsilon-1}\,dt \\ &= x^{-\alpha}\cdot\frac{x^\varepsilon-1}{\varepsilon} < \frac{x^{\varepsilon-\alpha}}{\varepsilon}, \end{aligned}$$

de lo que se deduce (45). Para llegar a (45) podríamos haber utilizado, también, el Teorema 8.6(f).

FUNCIONES TRIGONOMÉTRICAS

Definamos

$$C(x) = \frac{1}{2}[E(ix) + E(-ix)], \qquad S(x) = \frac{1}{2i}[E(ix) - E(-ix)]. \tag{46}$$

Demostraremos que $C(x)$ y $S(x)$ coinciden con las funciones $\cos x$ y $\operatorname{sen} x$, cuya definición se basa generalmente en consideraciones geométricas. Por (25), $E(\bar{z}) = \overline{E(z)}$. Por tanto, (46) demuestra que $C(x)$ y $S(x)$ son reales para x real. También

$$E(ix) = C(x) + iS(x). \tag{47}$$

Así pues, $C(x)$ y $S(x)$ son las partes real e imaginaria, respectivamente de $E(ix)$, si x es real. Por (27)

$$|E(ix)|^2 = E(ix)\overline{E(ix)} = E(ix)E(-ix) = 1,$$

de modo que

$$|E(ix)| = 1 \qquad (x \text{ real}). \tag{48}$$

De (46) podemos deducir que $C(0) = 1$; $S(0) = 0$ y (28) demuestra que

(49) $$C'(x) = -S(x), \qquad S'(x) = C(x).$$

Decimos que existen números positivos x para los cuales $C(x) = 0$. Supongamos que no es cierto. Como $C(0) = 1$, se deduce que $C(x) > 0$ para todo $x > 0$, por lo que $S'(x) > 0$, por (49) y S es estrictamente creciente; y como $S(0) = 0$ tenemos que $S(x) > 0$ si $x > 0$. Por tanto, si $0 < x < y$, tenemos

(50) $$S(x)(y-x) < \int_x^y S(t)\,dt = C(x) - C(y) \le 2.$$

La última desigualdad se deduce de (48) y (47). Como $S(x) > 0$, (50) no puede ser cierta para y grande, y tenemos una contradicción.

Sea x_0 el menor número positivo tal que $C(x_0) = 0$. Existe, pues el conjunto de ceros de una función continua es cerrado, y $C(0) \neq 0$. Definiremos el número π por

(51) $$\pi = 2x_0.$$

Entonces $C(\pi/2) = 0$, y (48) demuestra que $S(\pi/2) = \pm 1$. Como $C(x) > 0$ en $(0,\pi/2)$, S es creciente en $(0,\pi/2)$; por lo cual $S(\pi/2) = 1$. Así pues,

$$E\left(\frac{\pi i}{2}\right) = i,$$

y la fórmula de la adición da

(52) $$E(\pi i) = -1, \qquad E(2\pi i) = 1;$$

y de aquí

(53) $$E(z + 2\pi i) = E(z) \qquad (z \text{ compleja}).$$

8.7 Teorema

(*a*) *La función E es periódica, con periodo $2\pi i$.*
(*b*) *Las funciones C y S son periódicas, con periodo 2π.*
(*c*) *Si $0 < t < +2\pi$, $E(it) \neq 1$.*
(*d*) *Si z es un número complejo con $|z| = 1$, hay un único t en $[0,2\pi)$, tal que $E(it) = z$.*

Demostración Por (53), se cumple (*a*), y (*b*) se deduce de (*a*) y (46).

Supongamos que $0 < t < \pi/2$ y $E(it) = x + iy$, con x,y reales. De lo que hemos visto anteriormente se deduce que $0 < x < 1$, $0 < y < 1$. Obsérvese que

$$E(4it) = (x + iy)^4 = x^4 - 6x^2y^2 + y^4 + 4ixy(x^2 - y^2).$$

Si $E(4it)$ es real, se deduce que $x^2 - y^2 = 0$; como $x^2 + y^2 = 1$, por (48), tenemos que $x^2 = y^2 = \frac{1}{2}$, de donde $E(4it) = -1$, lo que demuestra (c).

Si $0 \leq t_1 < t_2 < 2\pi$, entonces de (c) se tiene

$$E(it_2)[E(it_1)]^{-1} = E(it_2 - it_1) \neq 1,$$

Esto establece la unicidad de (d).

Para demostrar la existencia en (d), se fija z de tal manera que $|z| = 1$, y se escribe $z = x + iy$, con x y y reales. Se supone enseguida que $x \geq 0$ y $y \geq 0$. Sobre $[0, \pi/2]$, C decrece de 1 a 0. De aquí que $C(t) = x$ para algún $t \in [0, \pi/2]$. Como $C^2 + S^2 = 1$ y $S \geq 0$ sobre $[0, \pi/2]$, se deduce que $z = E(it)$.

Si $x < 0$ y $y \geq 0$, entonces $-iz$ satisface las condiciones anteriores. En consecuencia $-iz = E(it)$ para algún $t \in [0, \pi/2]$, y como $i = E(\pi i/2)$, se obtiene $z = E(i(t + \pi/2))$. Para finalizar, si $y < 0$, los dos casos anteriores nos muestran que $-z = E(it)$ para algún $t \in (0, \pi)$. Por tanto $z = -E(it) = E(i(t + \pi))$.

Esto demuestra (d), y por ende el teorema.

De (d) y (48) se deduce que la curva y definida por

(54) $$\gamma(t) = E(it) \qquad (0 \leq t \leq 2\pi)$$

es una curva simple cerrada, cuyo rango es el círculo unidad en el plano. Como $\gamma'(t) = iE(it)$, la longitud de γ es

$$\int_0^{2\pi} |\gamma'(t)|\, dt = 2\pi,$$

por el Teorema 6.27. Este es, ciertamente, el resultado esperado para la circunferencia de un círculo de radio 1; y se ha demostrado que π, definido por (51), tiene el significado geómetrico usual.

Del mismo modo vemos que si t crece de 0 a t_0, el punto $\gamma(t)$ describe un arco circular de longitud t_0. Considerando el triángulo cuyos vértices son

$$z_1 = 0, \qquad z_2 = \gamma(t_0), \qquad z_3 = C(t_0)$$

se ve que $C(t)$ y $S(t)$ son realmente idénticas a cos t y sen t, si las últimas están definidas del modo habitual, por las razones entre los lados de un triángulo rectángulo.

Se recalca que hemos deducido las propiedades fundamentales de las funciones trigonométricas, a partir de (46) y (25) sin recurrir nunca a la noción geométrica de ángulo. Hay otros planteamientos no geométricos para estas funciones. Los siguientes artículos están relacionados con estos temas, el de W. F. Eberlein (*Amer. Math. Monthly,* tomo 74, 1967, páginas 1223 a 1225) y el de G. B. Robinson (*Math. Mag.* tomo 41, 1968, páginas 66 a 70).

LA COMPLETITUD ALGEBRAICA DEL CAMPO COMPLEJO

Ahora estamos en disposición de dar una demostración sencilla de la propiedad del campo complejo, de ser algebraicamente completo, es decir, que todo polinomio no constante con coeficientes complejos tiene una raíz compleja.

8.8 Teorema *Supongamos que $a_0, \ldots, a_n$ son números complejos, $n \geq 1$; $a_n \neq 0$;*

$$P(z) = \sum_0^n a_k z^k.$$

Entonces, $P(z) = 0$ para algún número complejo z.

Demostración Sin pérdida de generalidad, hagamos $a_n = 1$. Llamemos

$$\mu = \inf |P(z)| \qquad (z \text{ compleja}). \tag{55}$$

Si $|z| = R$,

$$|P(z)| \geq R^n[1 - |a_{n-1}|R^{-1} - \cdots - |a_0|R^{-n}]. \tag{56}$$

El segundo miembro de (56) tiende a ∞ cuando $R \to \infty$. Por tanto, existe R_0 tal que $|P(z)| > \mu$ si $|z| > R_0$. Como $|P|$ es continua en el disco cerrado con centro en 0 y radio R_0, el Teorema 4.16 demuestra que $|P(z_0)| = \mu$ para algún z_0.

Decimos que $\mu = 0$.

Si no lo es, hagamos $Q(z) = P(z + z_0)/P(z_0)$. Q será un polinomio no constante, $Q(0) = 1$, y $|Q(z)| \geq 1$ para todo z. Hay un menor entero, k, $1 \leq k \leq n$ tal que

$$Q(z) = 1 + b_k z^k + \cdots + b_n z^n, \qquad b_k \neq 0. \tag{57}$$

Por el Teorema 8.7(d) hay un θ real tal que

$$e^{ik\theta} b_k = -|b_k|. \tag{58}$$

Si $r > 0$ y $r^k |b_k| < 1$, (58) implica

$$|1 + b_k r^k e^{ik\theta}| = 1 - r^k |b_k|,$$

de modo que

$$|Q(re^{i\theta})| \leq 1 - r^k\{|b_k| - r|b_{k+1}| - \cdots - r^{n-k}|b_n|\}.$$

Para r suficientemente pequeño, la expresión entre llaves es positiva, por lo que $|Q(re^{i\theta})| < 1$, que es una contradicción.

Por tanto, $\mu = 0$, esto es, $P(z_0) = 0$.

El Ejercicio 27 contiene un resultado más general.

SERIES DE FOURIER

8.9 Definición Un *polinomio trigonométrico* es una suma finita de la forma

$$f(x) = a_0 + \sum_{n=1}^{N} (a_n \cos nx + b_n \operatorname{sen} nx) \qquad (x \text{ real}), \tag{59}$$

donde $a_0, \ldots, a_N, b_1, \ldots, b_N$ son números complejos. Teniendo en cuenta las identidades (46), también puede escribirse (59) en la forma

$$f(x) = \sum_{-N}^{N} c_n e^{inx} \qquad (x \text{ real}), \tag{60}$$

que es más conveniente en muchos casos. Es claro que todo polinomio trigonométrico es periódico, con periodo 2π.

Si n es un entero no nulo, e^{inx} es la derivada de e^{inx}/in, que tiene también periodo 2π. Por tanto,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\,dx = \begin{cases} 1 & (\text{si } n = 0), \\ 0 & (\text{si } n = \pm 1, \pm 2, \ldots). \end{cases} \tag{61}$$

Multipliquemos (60) por e^{-imx}, donde m es entero; si integramos el producto, (61) demuestra que

$$c_m = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-imx}\,dx \tag{62}$$

para $|m| \leq N$. Si $|m| > N$, la integral de (62) es 0.

La observación siguiente puede deducirse de (60) y (62): El polinomio trigonométrico f, dado por (60), es *real* si y solo si $c_{-n} = \overline{c_n}$ para $n = 0, \ldots, N$.

De acuerdo con (60), definiremos como *serie trigonométrica* a la serie de la forma

$$\sum_{-\infty}^{\infty} c_n e^{inx} \qquad (x \text{ real}); \tag{63}$$

definida la suma parcial N-ésima de (63) como el segundo miembro de (60).

Si f es una función integrable en $[-\pi,\pi]$, los números c_m definidos por (62) para todo entero m se llaman *coeficientes de Fourier* de f, y la serie (63), formada con estos coeficientes, es la *serie de Fourier* de f.

La pregunta principal que ahora se plantea, es si la serie de Fourier de f converge hacia f o, más general, si f está determinada por su serie de Fourier. Es decir, si conocemos los coeficientes de Fourier de una función, ¿podemos hallar la función?, y si es así, ¿cómo?

El estudio de tales series, y, en particular el problema de representar una función dada por una serie trigonométrica, se originó en los problemas de física, tales como la teoría de oscilaciones y la de la transmisión del calor. («La Théorie analytique de la chaleur» de Fourier se publicó en 1822.) Las muchas dificultades y delicados problemas aparecidos durante este estudio originaron una revisión y reformulación de toda la teoría de las funciones de variable real. Entre muchos nombres famosos, los de Reimann, Cantor y Lebesgue están íntimamente ligados con este tema, que en la actualidad, con todas sus generalizaciones y ramificaciones, puede decirse que ocupa el puesto central de todo el análisis.

Vamos, ahora, a deducir algunos teoremas fundamentales que están fácilmente a nuestro alcance con los métodos desarrollados en los capítulos precedentes. Para un estudio más profundo, la integral de Lebesgue es un instrumento natural e indispensable.

Primeramente estudiaremos sistemas de funciones más generales que participan de propiedades análogas a (61):

8.10 Definición Sea $\{\phi_n\}$ $(n = 1, 2, 3, \ldots)$ una sucesión de funciones complejas en $[a,b]$, tales que

$$\int_a^b \phi_n(x)\overline{\phi_m(x)}\, dx = 0 \qquad (n \neq m). \tag{64}$$

Entonces, se dice que $\{\phi_n\}$ es un *sistema ortogonal de funciones* en $[a,b]$. Si, además,

$$\int_a^b |\phi_n(x)|^2\, dx = 1 \tag{65}$$

para todo n, se dice que $\{\phi_n\}$ es *ortonormal u ortogonal y normalizada*.

Por ejemplo, las funciones $(2\pi)^{-\frac{1}{2}}e^{inx}$ forman un sistema ortonormal en $[-\pi,\pi]$, lo mismo que las funciones reales

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\operatorname{sen} x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\operatorname{sen} 2x}{\sqrt{\pi}}, \dots .$$

Si $\{\phi_n\}$ es ortonormal en $[a,b]$ y si

(66) $$c_n = \int_a^b f(t)\overline{\phi_n(t)}\,dt \qquad (n = 1, 2, 3, \dots),$$

llamaremos a c_n coeficiente n-ésimo de Fourier de f relativo a $\{\phi_n\}$. Escribiremos

(67) $$f(x) \sim \sum_1^{\infty} c_n \phi_n(x)$$

y llamaremos a esta serie, serie de Fourier de f (relativa a $\{\phi_n\}$).

Obsérvese que el símbolo $\sim$ utilizado en (67) no implica nada sobre la convergencia de la serie; sino que expresa solamente que los coeficientes vienen dados por (66).

Los siguientes teoremas demuestran que las sumas parciales de la serie de Fourier de f tienen una cierta propiedad mínima. Admitiremos aquí, y en lo que resta de capítulo, que $f \in \mathscr{R}$, aunque se puede hacer una hipótesis más amplia.

8.11 Teorema *Sea $\{\phi_n\}$ ortonormal en $[a,b]$. Sea*

(68) $$s_n(x) = \sum_{m=1}^{n} c_m \phi_m(x)$$

la n-ésima suma parcial de la serie de Fourier de f, y supongamos que

(69) $$t_n(x) = \sum_{m=1}^{n} \gamma_m \phi_m(x).$$

Será

(70) $$\int_a^b |f - s_n|^2\,dx \le \int_a^b |f - t_n|^2\,dx,$$

y la igualdad se cumple si, y solo si

(71) $$\gamma_m = c_m \qquad (m = 1, \dots, n).$$

Es decir, entre todas las funciones t_n, s_n da la mejor aproximación cuadrática media de f.

Demostración Sea $\int$ la integral en $[a,b]$, Σ la suma de 1 a n. Entonces,

$$\int f\bar{t}_n = \int f \sum \bar{\gamma}_m \bar{\phi}_m = \sum c_m \bar{\gamma}_m$$

por la definición de $\{c_m\}$,

$$\int |t_n|^2 = \int t_n \bar{t}_n = \int \sum \gamma_m \phi_m \sum \bar{\gamma}_k \bar{\phi}_k = \sum |\gamma_m|^2$$

puesto que $\{\phi_m\}$ es ortonormal, y así,

$$\begin{aligned}\int |f - t_n|^2 &= \int |f|^2 - \int f\bar{t}_n - \int \bar{f} t_n + \int |t_n|^2 \\ &= \int |f|^2 - \sum c_m \bar{\gamma}_m - \sum \bar{c}_m \gamma_m + \sum \gamma_m \bar{\gamma}_m \\ &= \int |f|^2 - \sum |c_m|^2 + \sum |\gamma_m - c_m|^2,\end{aligned}$$

qué, evidentemente es mínima si, y solo si $\gamma_m = c_m$.

Poniendo $\gamma_m = c_m$ es este cálculo, obtenemos

(72) $$\int_a^b |s_n(x)|^2\, dx = \sum_1^n |c_m|^2 \leq \int_a^b |f(x)|^2\, dx,$$

pues $\int |f - t_n|^2 \geq 0$.

8.12 Teorema *Si $\{\phi_n\}$ es ortonormal en $[a,b]$, y si*

$$f(x) \sim \sum_{n=1}^{\infty} c_n \phi_n(x),$$

será

(73) $$\sum_{n=1}^{\infty} |c_n|^2 \leq \int_a^b |f(x)|^2\, dx.$$

En particular,

(74) $$\lim_{n\to\infty} c_n = 0.$$

Demostración Haciendo $n \to \infty$ en (72), obtenemos (73), llamada «desigualdad de Bessel».

8.13 Serie Trigonométrica De ahora en adelante, se tratará sólo con el sistema trigonométrico. Considerando funciones f que tienen periodo 2π y son

integrables sobre $[-\pi,\pi]$ según Riemann (y por esto son acotadas en cada intervalo). La serie de Fourier de f es la (63) cuyos coeficientes c_n se calculan por medio de (62), y

$$s_N(x) = s_N(f;\, x) = \sum_{-N}^{N} c_n e^{inx} \tag{75}$$

es la N-ésima suma parcial de la serie de Fourier de f. La desigualdad (72) toma ahora la forma

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |s_N(x)|^2\, dx = \sum_{-N}^{N} |c_n|^2 \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\, dx. \tag{76}$$

Se introduce ahora el *núcleo de Dirichlet*

$$D_N(x) = \sum_{n=-N}^{N} e^{inx} = \frac{\operatorname{sen}(N+\frac{1}{2})x}{\operatorname{sen}(x/2)}. \tag{77}$$

que servirá para obtener una expresión para S_N que es más fácil de ma. jar que la (75). La primera de las igualdades en (77) es la definición de $D_N(x)$. La segunda se deduce si ambos lados de la identidad

$$(e^{ix} - 1)D_N(x) = e^{i(N+1)x} - e^{-iNx}$$

se multiplica por $e^{-ix/2}$.

De (62) y (75), se tiene

$$\begin{aligned} s_N(f;\, x) &= \sum_{-N}^{N} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\, dt\; e^{inx} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) \sum_{-N}^{N} e^{in(x-t)}\, dt, \end{aligned}$$

de modo que

$$s_N(f;\, x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_N(x-t)\, dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)\, dt. \tag{78}$$

La periodicidad de todas las funciones involucradas muestra que no importa sobre que intervalo se integre, mientras su longitud sea 2π. Esto muestra que las dos integrales de (78) son iguales.

Se probará solo un teorema sobre la convergencia puntual de las series de Fourier.

8.14 Teorema *Si para algún x, hay constantes* $\delta > 0$ *y* $M < \infty$ *tales que*

(79) $$|f(x+t)-f(x)| \le M|t|$$

para todo $t \in (-\delta,\delta)$, *entonces*

(80) $$\lim_{N\to\infty} s_N(f;x) = f(x).$$

Demostración Defínase

(81) $$g(t) = \frac{f(x-t)-f(x)}{\operatorname{sen}(t/2)}$$

para $0 < |t| \le \pi$, y hágase $g(0) = 0$. De la definición (77),

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(x)\,dx = 1.$$

De esto, (78) muestra que

$$s_N(f;x) - f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\operatorname{sen}\left(N+\frac{1}{2}\right)t\,dt$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[g(t)\cos\frac{t}{2}\right]\operatorname{sen} Nt\,dt + \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[g(t)\operatorname{sen}\frac{t}{2}\right]\cos Nt\,dt.$$

De (79) y (81) se ve que $g(t)\cos(t/2)$ y $g(t)\operatorname{sen}(t/2)$ son acotadas. Entonces las últimas dos integrales tienden hacia 0 cuando $N\to\infty$, por (74). Esto demuestra (80).

Corolario *Si* $f(x) = 0$ *para todo* x *en algún segmento* J, *entonces* lím $s_N(f;x) = 0$ *para cada* $x \in J$.

He aquí otra formulación de este corolario:
Si $f(t) = g(t)$ *para todo* t *en alguna vecindad de* x, *entonces*

$$s_N(f;x) - s_N(g;x) = s_N(f-g;x) \to 0 \text{ como } N\to\infty.$$

Este se denomina comúnmente *teorema de localización*. Muestra que el comportamiento de la sucesión $\{s_N(f;x)\}$, por lo que a convergencia se refiere, depende sólo de los valores de f en alguna vecindad de x (arbitrariamente pequeña). Entonces dos series de Fourier pueden tener el mismo comportamiento en un intervalo, pero pueden comportarse de manera totalmente diferente en otro intervalo. Aquí se tiene un contraste muy notorio entre las series de Fourier y las de potencias (Teorema 8.5).

Se concluirá esta sección con otros dos teoremas de aproximación.

8.15 Teorema *Si f es continua (con periodo 2π) y si $\varepsilon > 0$, entonces hay un polinomio trigonométrico P tal que*

$$|P(x) - f(x)| < \varepsilon$$

para todo real x.

Demostración Si se identifican x y $x + 2\pi$, se puede considerar a las funciones con periodo 2π sobre R^1 como funciones sobre el círculo unitario T como el mapeo $x \to e^{ix}$. Los polinomios trigonométricos, es decir, las funciones de la forma (60), forman un álgebra autoadjunta $\mathscr{A}$, que separa puntos sobre T, y que no desaparece en ningún punto de T. Como T es compacto, el Teorema 7.33 indica que $\mathscr{A}$ es denso en $\mathscr{C}(T)$. Esto es lo que con exactitud asegura el teorema.

Una forma más precisa de este teorema se presenta en el Ejercicio 15.

8.16 Teorema de Parseval *Supóngase que f y g son funciones integrables según Riemann y que tienen periodo 2π, y además*

$$f(x) \sim \sum_{-\infty}^{\infty} c_n e^{inx}, \qquad g(x) \sim \sum_{-\infty}^{\infty} \gamma_n e^{inx}. \tag{82}$$

Entonces

$$\lim_{N\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - s_N(f; x)|^2\, dx = 0, \tag{83}$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\, dx = \sum_{-\infty}^{\infty} c_n \bar{\gamma}_n, \tag{84}$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\, dx = \sum_{-\infty}^{\infty} |c_n|^2. \tag{85}$$

Demostración Usemos la notación

$$\|h\|_2 = \left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} |h(x)|^2\, dx\right\}^{1/2}. \tag{86}$$

Dado $\varepsilon > 0$, como $f \in \mathscr{R}$ y $f(\pi) = f(-\pi)$, la construcción que se describió en el Ejercicio 12 del capítulo 6 proporciona una función continua 2π con periodo h tal que

$$\|f - h\|_2 < \varepsilon. \tag{87}$$

Del Teorema 8.15, se puede ver que hay un polinomio trigonométrico P tal que $|h(x) - P(x)| < \varepsilon$ para todo x. De aquí que $\|h - P\|_2 < \varepsilon$. Si P es de grado N_0, el Teorema 8.11 muestra que

(88) $$\|h - s_N(h)\|_2 \leq \|h - P\|_2 < \varepsilon$$

Para todo $N \geq N_0$. De (72), y con $h - f$ en vez de f,

(89) $$\|s_N(h) - s_N(f)\|_2 = \|s_N(h - f)\|_2 \leq \|h - f\|_2 < \varepsilon.$$

Ahora, la desigualdad del triángulo (Ejercicio 11, Cap. 6) combinada con (87), (88) y (89), muestra que

(90) $$\|f - s_N(f)\|_2 < 3\varepsilon \qquad (N \geq N_0).$$

Esto demuestra (83). A continuación,

(91) $$\frac{1}{2\pi}\int_{-\pi}^{\pi} s_N(f)\bar{g}\,dx = \sum_{-N}^{N} c_n \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\overline{g(x)}\,dx = \sum_{-N}^{N} c_n \bar{\gamma}_n,$$

y la desigualdad de Schwarz muestra que

(92) $$\left|\int f\bar{g} - \int s_N(f)\bar{g}\right| \leq \int |f - s_N(f)||g| \leq \left\{\int |f - s_N|^2 \int |g|^2\right\}^{1/2},$$

la cual tiende hacia 0, cuando $N \to \infty$, por (83). Comparando (91) y (92) se obtiene (84). Finalmente, de (84) se ve que (85) es el caso especial $g = f$.

En el capítulo 11 se presenta una versión más general del Teorema 8.16.

LA FUNCIÓN GAMMA

Esta función está muy relacionada con los factoriales y aparece inesperadamente en el análisis. En el artículo de P. J. Davis (*Amer. Math. Monthly*, tomo 66, 1959, páginas 849 a la 869), se describen muy bien el origen, la historia y el desarrollo de esta función. Otra introducción elemental bastante buena, es el libro de Artin (que se cita en la Bibliografía).

La presentación será bastante condensada, y sólo se harán algunos comentarios después de cada teorema. Se puede considerar esta sección como un ejercicio bastante largo, y como oportunidad para aplicar algo del material presentado hasta ahora.

8.17 Definición Para $0 < x < \infty$,

(93) $$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt.$$

La integral converge para estas x. (Cuando $x < 1$, tienen que observarse 0 y ∞.)

8.18 Teorema

(*a*) *La ecuación funcional*

$$\Gamma(x+1) = x\Gamma(x)$$

se cumple si $0 < x < \infty$.
(*b*) $\Gamma(n+1) = n!$ *para* $n = 1, 2, 3, \ldots$
(*c*) $\log \Gamma$ *es convexa sobre* $(0, \infty)$.

Demostración Integrando por partes se demuestra (*a*). Como $\Gamma(1) = 1$, por inducción (*a*) implica (*b*). Si $1 < p < \infty$ y $(1/p) + (1/q) = 1$, se aplica la desigualdad de Hölder (Ejercicio 10, Cap. 6) a (93), y se obtiene

$$\Gamma\left(\frac{x}{p} + \frac{y}{q}\right) \le \Gamma(x)^{1/p}\Gamma(y)^{1/q}.$$

Esto es equivalente a (*c*).

Un descubrimiento sorprendente de Bohr y Mollerup es que estas tres propiedades caracterizan completamente a Γ.

8.19 Teorema *Si f es una función positiva sobre* $(0, \infty)$ *tal que*

(*a*) $f(x+1) = xf(x)$,
(*b*) $f(1) = 1$,
(*c*) $\log f$ *es convexa*,
entonces $f(x) = \Gamma(x)$.

Demostración Como Γ satisface (*a*), (*b*) y (*c*), es suficiente demostrar que $f(x)$ se determina de manera única por (*a*), (*b*), (*c*), para todo $x > 0$. Por (*a*), basta con hacerlo para $x \in (0, 1)$.

Si se hace $\varphi = \log f$. Entonces

(94) $$\varphi(x+1) = \varphi(x) + \log x \qquad (0 < x < \infty),$$

$\varphi(1) = 0$, y φ es convexa. Supóngase que $0 < x < 1$, y n es un entero positivo. De (94), $\varphi(n + 1) = \log(n!)$. Considérese ahora el cociente de diferencias de φ sobre los intervalos $[n, n + 1]$, $[n + 1, n + 1 + x]$, $[n + 1, n + 2]$. Como φ es convexa,

$$\log n \leq \frac{\varphi(n + 1 + x) - \varphi(n + 1)}{x} \leq \log (n + 1).$$

La aplicación repetida de (94) produce

$$\varphi(n + 1 + x) = \varphi(x) + \log [x(x + 1) \cdots (x + n)].$$

Entonces

$$0 \leq \varphi(x) - \log \left[\frac{n!n^x}{x(x + 1) \cdots (x + n)}\right] \leq x \log \left(1 + \frac{1}{n}\right).$$

La última expresión tiende hacia 0 cuando $n \to \infty$. De aquí que se determine $\varphi(x)$, y esto completa la demostración.

La relación siguiente se obtiene por un producto

(95) $$\Gamma(x) = \lim_{n \to \infty} \frac{n!n^x}{x(x + 1) \cdots (x + n)}$$

al menos cuando $0 < x < 1$; de esto se puede deducir que (95) es cierta para todo $x > 0$, ya que $\Gamma(x + 1) = x\Gamma(x)$.

8.20 Teorema *Si $x > 0$ y $y > 0$, entonces*

(96) $$\int_0^1 t^{x-1}(1 - t)^{y-1}\, dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x + y)}.$$

Esta integral es la llamada *función beta $B(x,y)$*.

Demostración Notar que $B(1,y) = 1/y$, y que $\log B(x,y)$ es una función convexa de la variable x, para cada y fijo, por la desigualdad de Hölder, de la misma forma que en el Teorema 8.18, y además,

(97) $$B(x + 1, y) = \frac{x}{x + y} B(x, y).$$

Para demostrar (97), se efectúa una integración por partes sobre

$$B(x+1,y)=\int_0^1 \left(\frac{t}{1-t}\right)^x (1-t)^{x+y-1}\,dt.$$

Estas tres propiedades de $B(x,y)$ muestran para cada y, que el Teorema 8.19 se aplica a la función f definida por

$$f(x)=\frac{\Gamma(x+y)}{\Gamma(y)}B(x,y).$$

En consecuencia, $f(x) = \Gamma(x)$.

8.21 Algunas consecuencias La sustitución $t = \operatorname{sen}^2\theta$ convierte a (96) en

(98) $$2\int_0^{\pi/2} (\operatorname{sen}\theta)^{2x-1}(\cos\theta)^{2y-1}\,d\theta=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

El caso especial $x = y = \frac{1}{2}$ produce

(99) $$\Gamma(\tfrac{1}{2})=\sqrt{\pi}.$$

La sustitución $t = s^2$ convierte a (93) en

(100) $$\Gamma(x)=2\int_0^\infty s^{2x-1}e^{-s^2}\,ds \qquad (0<x<\infty).$$

El caso especial $x = \frac{1}{2}$ produce

(101) $$\int_{-\infty}^\infty e^{-s^2}\,ds=\sqrt{\pi}.$$

Por (99), la identidad

(102) $$\Gamma(x)=\frac{2^{x-1}}{\sqrt{\pi}}\Gamma\left(\frac{x}{2}\right)\Gamma\left(\frac{x+1}{2}\right)$$

se deduce directamente del Teorema 8.19.

8.22 Fórmula de Stirling Esta proporciona una expresión aproximada sencilla para $\Gamma(x+1)$ cuando x es grande (por consiguiente para $n!$ cuando n es grande). La fórmula es

(103) $$\lim_{x\to\infty}\frac{\Gamma(x+1)}{(x/e)^x\sqrt{2\pi x}}=1.$$

He aquí una demostración. Haciendo $t = x(1+u)$ en (93). Esto da

$$\Gamma(x+1) = x^{x+1} e^{-x} \int_{-1}^{\infty} [(1+u)e^{-u}]^x \, du. \tag{104}$$

Se determina $h(u)$ de manera que $h(0) = 1$ y

$$(1+u)e^{-u} = \exp\left[-\frac{u^2}{2} h(u)\right] \tag{105}$$

Si $-1 < u < \infty$, $u \neq 0$. Entonces

$$h(u) = \frac{2}{u^2}[u - \log(1+u)]. \tag{106}$$

Y se deduce que h es continua, y que $h(u)$ decrece monótonamente desde ∞ hasta 0 conforme u crece desde -1 hasta ∞.

La sustitución $u = s\sqrt{2/x}$ convierte a (104) en

$$\Gamma(x+1) = x^x e^{-x} \sqrt{2x} \int_{-\infty}^{\infty} \psi_x(s) \, ds \tag{107}$$

en donde

$$\psi_x(s) = \begin{cases} \exp[-s^2 h(s\sqrt{2/x})] & (-\sqrt{x/2} < s < \infty), \\ 0 & (s \leq -\sqrt{x/2}). \end{cases}$$

Nótense las siguientes propiedades de $\psi_x(s)$:

(*a*) Para cada s, $\psi_x(s) \to e^{-s^2}$ cuando $x \to \infty$.
(*b*) La convergencia en (*a*) es uniforme sobre $[-A, A]$, para cada $A < \infty$.
(*c*) Cuando $s < 0$, entonces $0 < \psi_x(s) < e^{-s^2}$.
(*d*) Cuando $s > 0$ y $x > 1$, entonces $0 < \psi_x(s) < \psi_1(s)$.
(*e*) $\int_0^\infty \psi_1(s) \, ds < \infty$.

El teorema de convergencia que se estableció en el Ejercicio 12 del capítulo 7, puede aplicarse a la integral (107), y muestra que ésta converge hacia $\sqrt{\pi}$ cuando $x \to \infty$, por (101). Esto demuestra (103).

En el libro "Advanced Calculus" de R. C. Buck hay una versión más detallada de esta demostración (páginas 216 a 218.) Para otras demostraciones completamente diferentes, véanse el artículo de W. Feller en el *Amer. Math. Monthly, tomo 74, 1967, páginas 1223 a 1225 (con una corrección en* el tomo 75, 1968, página 518), y las páginas 20 a 24 del libro de Artin.

El Ejercicio 20 da una demostración simple de un resultado más sencillo.

EJERCICIOS

1. Definida

$$f(x)=\begin{cases} e^{-1/x^2} & (x\neq 0),\\ 0 & (x=0).\end{cases}$$

Demostrar que f tiene derivadas de todos los órdenes en $x = 0$, y que $f^{(n)}(0) = 0$, para $n = 1, 2, 3, \ldots$

2. Sea a_{ij} el número que ocupa la fila i y la columna j en el cuadro

$$\begin{array}{ccccc} -1 & 0 & 0 & 0 & \cdots \\ \frac{1}{2} & -1 & 0 & 0 & \cdots \\ \frac{1}{4} & \frac{1}{2} & -1 & 0 & \cdots \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{2} & -1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{array}$$

tal que

$$a_{ij}=\begin{cases} 0 & (i<j),\\ -1 & (i=j),\\ 2^{j-i} & (i>j).\end{cases}$$

Probar que

$$\sum_i \sum_j a_{ij} = -2, \qquad \sum_j \sum_i a_{ij} = 0.$$

3. Demostrar que

$$\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij}$$

si $a_{ij} \geq 0$ para todo i y j (puede presentarse el caso $+\infty = +\infty$).

4. Demostrar las siguientes expresiones sobre límites:

(*a*) $\displaystyle \lim_{x\to 0} \frac{b^x-1}{x} = \log b \qquad (b>0).$

(*b*) $\displaystyle \lim_{x\to 0} \frac{\log(1+x)}{x} = 1.$

(*c*) $\displaystyle \lim_{x\to 0} (1+x)^{1/x} = e.$

(*d*) $\displaystyle \lim_{n\to\infty} \left(1+\frac{x}{n}\right)^n = e^x.$

5. Encontrar los límites siguientes:

(a) $\lim\limits_{x\to 0} \dfrac{e-(1+x)^{1/x}}{x}$.

(b) $\lim\limits_{n\to \infty} \dfrac{n}{\log n}[n^{1/n}-1]$.

(c) $\lim\limits_{x\to 0} \dfrac{\tan x - x}{x(1-\cos x)}$.

(d) $\lim\limits_{x\to 0} \dfrac{x-\operatorname{sen} x}{\tan x - x}$.

6. Suponer $f(x)f(y) = f(x+y)$ para todo x y y reales.
(a) Admitiendo que f es diferenciable y no nula, probar que

$$f(x)=e^{cx}$$

siendo c una constante.
(b) Demostrar lo mismo, suponiendo solamente que f es continua.

7. Si $0<x<\dfrac{\pi}{2}$, probar que

$$\frac{2}{\pi}<\frac{\operatorname{sen} x}{x}<1.$$

8. Para $n = 0, 1, 2, \ldots,$ y x real, probar que

$$|\operatorname{sen} nx| \le n|\operatorname{sen} x|.$$

Obsérvese que esta desigualdad puede ser falsa para otros valores de n. Por ejemplo,

$$|\operatorname{sen} \tfrac{1}{2}\pi| > \tfrac{1}{2}|\operatorname{sen} \pi|.$$

9. (a) Haciendo $s_N = 1 + (\frac{1}{2}) + \cdots + (1/N)$ Demostrar que

$$\lim_{N\to\infty} (s_N - \log N)$$

existe. (Este límite se representa comúnmente por γ, y se le llama constante de Euler. Su valor numérico es 0.5772..., y no se sabe si γ es racional o no.)
(b) Aproximadamente, ¿qué tan grande debe ser m para que $N = 10^m$ satisfaga la desigualdad $s_N > 100$?

10. Demostrar que $\Sigma\ 1/p$ diverge; la suma se efectúa sobre todos los primos.

(Esto demuestra que los primos forman un subconjunto bastante substancial de los enteros positivos.)

Sugerencia: Dado N, sean $p_1, \ldots, p_k$ los primos que dividen al menos a un entero $\le N$. Entonces

$$\sum_{n=1}^{N} \frac{1}{n} \leq \prod_{j=1}^{k} \left(1 + \frac{1}{p_j} + \frac{1}{p_j^2} + \cdots\right)$$
$$= \prod_{j=1}^{k} \left(1 - \frac{1}{p_j}\right)^{-1}$$
$$\leq \exp \sum_{j=1}^{k} \frac{2}{p_j}.$$

Esta última desigualdad se cumple porque

$$(1-x)^{-1} \leq e^{2x}$$

si $0 \leq x \leq \frac{1}{2}$.

(Existen muchas demostraciones de este resultado. Véanse, por ejemplo, el artículo de I. Niven en la *Amer. Math. Monthly*, tomo 78, 1971, páginas 272 a 273, y el de R. Bellman en la *Amer. Math. Monthly*, tomo 50, 1943, páginas 318 a 319.)

11. Supóngase que $f \in \mathscr{R}$ sobre $[0, A]$ para todo $A < \infty$, y $f(x) \to 1$ cuando $x \to +\infty$. Demostrar que

$$\lim_{t \to 0} t \int_0^\infty e^{-tx} f(x)\, dx = 1 \qquad (t > 0).$$

12. Supóngase que $0 < \delta < \pi$, $f(x) = 1$ si $|x| \leq \delta$, $f(x) = 0$, si $\delta < (x) \leq \pi$, y $f(x + 2\pi) = f(x)$ para todo x.

(*a*) Calcular los coeficientes de Fourier de f.

(*b*) Concluir que

$$\sum_{n=1}^{\infty} \frac{\operatorname{sen}(n\delta)}{n} = \frac{\pi - \delta}{2} \qquad (0 < \delta < \pi).$$

(*c*) A partir del teorema de Parseval, deducir que

$$\sum_{n=1}^{\infty} \frac{\operatorname{sen}^2(n\delta)}{n^2\delta} = \frac{\pi - \delta}{2}.$$

(*d*) Hacer $\delta \to 0$ y demostrar que

$$\int_0^\infty \left(\frac{\operatorname{sen} x}{x}\right)^2 dx = \frac{\pi}{2}.$$

(*e*) Poner en (*c*) $\delta = \pi/2$. ¿Qué se obtiene?

13. Hacer $f(x) = x$ si $0 \leq x < 2\pi$, y aplicar el teorema de Parseval para concluir que

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

14. Si $f(x) = (\pi - |x|)^2$ sobre $[-\pi, \pi]$, demuéstrese que

$$f(x) = \frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{4}{n^2} \cos nx$$

y dedúzcase que

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}, \qquad \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}.$$

(Un artículo reciente de E. L. Stark contiene muchas referencias sobre series de la forma Σn^{-s}, donde s es un entero positivo. Véase *Math. Mag.*, tomo 47, 1974, páginas 197 a 202.)

15. Con D_n como se definió en (77), y poniendo

$$K_N(x) = \frac{1}{N+1} \sum_{n=0}^{N} D_n(x).$$

Demostrar que

$$K_N(x) = \frac{1}{N+1} \cdot \frac{1 - \cos (N+1)x}{1 - \cos x}$$

y también que

(*a*) $K_N \geq 0$,

(*b*) $\frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(x)\, dx = 1$,

(*c*) $K_N(x) \leq \frac{1}{N+1} \cdot \frac{2}{1 - \cos \delta}$ Si $0 < \delta \leq |x| \leq \pi$.

Si $s_N = s_N(f;\, x)$ es la N-ésima suma parcial de la serie de Fourier de f, considérese la media aritmética

$$\sigma_N = \frac{s_0 + s_1 + \cdots + s_N}{N+1}.$$

Demostrar que

$$\sigma_N(f;\, x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) K_N(t)\, dt,$$

y por consiguiente probar el teorema de Fejér:

Si f es continua, con periodo 2π, entonces $\sigma_N(f;\, x) \to f(x)$ uniformemente sobre $[-\pi, \pi]$. Sugerencia: Usar las propiedades (*a*), (*b*), (*c*) y procédase como en el Teorema 7.26.

16. Demostrar la siguiente versión puntual del teorema de Fejér:

Si $f \in \mathscr{R}$ y $f(x+)$, $f(x-)$ existe para algún x, entonces

$$\lim_{N \to \infty} \sigma_N(f;\, x) = \tfrac{1}{2}[f(x+) + f(x-)].$$

17. Supóngase que f es acotada y monótona sobre $[-\pi, \pi]$, con coeficientes de Fourier c_n, como los dados en (62).

(*a*) Usar el Ejercicio 17 del capítulo 6 para demostrar que $\{nc_n\}$ es una sucesión acotada.

(*b*) Combinar (*a*) con el Ejercicio 16 y el Ejercicio 14(*e*) del capítulo 3 para poder concluir que

$$\lim_{N\to\infty} s_N(f;x) = \tfrac{1}{2}[f(x+)+f(x-)]$$

para cada x.

(*c*) Supóngase sólo que $f \in \mathscr{R}$ sobre $[-\pi,\pi]$ y que f es monótona en algún segmento $(\alpha,\beta) \subset [-\pi,\pi]$. Demostrar que la conclusión de (*b*) se cumple para todo $x \in (\alpha,\beta)$.

(Esta es una aplicación del teorema de localización.)

18. Se define

$$f(x) = x^3 - \operatorname{sen}^2 x \tan x$$

$$g(x) = 2x^2 - \operatorname{sen}^2 x - x \tan x.$$

Para cada una de estas dos funciones, determinar si son positivas o negativas para todo $x \in (0, \pi/2)$, o si cambian de signo. Demuestre su respuesta.

19. Supóngase que f es una función continua sobre R^1, $f(x + 2\pi) = f(x)$, y α/π es irracional. Demostrar que

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x+n\alpha) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt$$

para cada x. *Sugerencia*: Hacerlo primero para $f(x) = e^{ikx}$.

20. El cálculo sencillo que se muestra a continuación, proporciona una aproximación buena a la fórmula de Stirling.

Para $m = 1, 2, 3, \ldots$, se define

$$f(x) = (m+1-x)\log m + (x-m)\log(m+1)$$

si $m \le x \le m + 1$, y se define también

$$g(x) = \frac{x}{m} - 1 + \log m$$

si $m - \frac{1}{2} \le x < m + \frac{1}{2}$. Dibujar las gráficas de f y g. Notar que $f(x) \le \log x \le g(x)$ si $x \ge 1$ y que

$$\int_1^n f(x)\,dx = \log(n!) - \tfrac{1}{2}\log n > -\tfrac{1}{8} + \int_1^n g(x)\,dx.$$

Integrar $\log x$ sobre $[1, n]$. Concluir que

$$\tfrac{7}{8} < \log(n!) - (n+\tfrac{1}{2})\log n + n < 1$$

para $n = 2, 3, 4, \ldots$. (*Nota*: $\log\sqrt{2\pi} \sim 0.918\ldots$.) Entonces

$$e^{7/8} < \frac{n!}{(n/e)^n\sqrt{n}} < e.$$

21. Sea

$$L_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|\, dt \qquad (n = 1, 2, 3, \ldots).$$

Demostrar que existe una constante $C > 0$ tal que

$$L_n > C \log n \qquad (n = 1, 2, 3, \ldots),$$

o, dicho de manera más precisa, que la sucesión

$$\left\{L_n - \frac{4}{\pi^2}\log n\right\}$$

es acotada.

22. Si α es real, y $-1 < x < 1$, demostrar el teorema del binomio de Newton

$$(1+x)^\alpha = 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n.$$

Sugerencia: Representar el miembro derecho con $f(x)$. Demostrar que la serie converge. Demostrar después que

$$(1+x)f'(x) = \alpha f(x)$$

y resolver esta ecuación diferencial.

Mostrar también que

$$(1-x)^{-\alpha} = \sum_{n=0}^{\infty} \frac{\Gamma(n+\alpha)}{n!\,\Gamma(\alpha)}x^n$$

si $-1 < x < 1$ y $\alpha > 0$.

23. Sea γ una curva *cerrada* continuamente diferenciable en el plano complejo, con un intervalo de parámetros $[a,b]$, y admítase que $\gamma(t) \neq 0$ para cada $t \in [a,b]$. Se define el *índice* de γ como

$$\text{Ind}\,(\gamma) = \frac{1}{2\pi i}\int_a^b \frac{\gamma'(t)}{\gamma(t)}\, dt.$$

Demostrar que el Índ (γ) es siempre un entero.

Sugerencia: Existe φ sobre $[a,b]$ con $\varphi' = \gamma'/\gamma$, $\varphi(a) = 0$. En consecuencia $\gamma\exp(-\varphi)$ es constante. Ya que $\gamma(a) = \gamma(b)$, se deduce que $\exp\varphi(b) = \exp\varphi(a) = 1$. Notar que $\varphi(b) = 2\pi i$ Índ (γ).

Calcular Índ (γ) cuando $\gamma(t) = e^{int}$, $a = 0$, $b = 2\pi$.

Explicar por qué Índ (γ) se denomina comúnmente el *número de arrollamiento* de γ alrededor de 0.

24. Sea γ como en el Ejercicio 23, y supóngase además que el rango de γ no interseca el eje real negativo. Demostrar que Índ $(\gamma) = 0$. *Sugerencia*: Para $0 \le c < \infty$, Índ $(\gamma + c)$ es una función de c continua valuada en los enteros. También, Índ $(\gamma + c) \to 0$ cuando $c \to \infty$.

25. Supóngase que γ_1 y γ_2 son las curvas definidas como en el Ejemplo 23, y

$$|\gamma_1(t)-\gamma_2(t)|<|\gamma_1(t)| \qquad (a\leq t\leq b).$$

Demostrar que Índ (γ_1) = Índ (γ_2).

Sugerencia: Hágase $\gamma = \gamma_2/\gamma_1$, Enseguida $|1-\gamma|<1$, de aquí que por el Ejercicio 24, Índ $(\gamma) = 0$. También

$$\frac{\gamma'}{\gamma}=\frac{\gamma_2'}{\gamma_2}-\frac{\gamma_1'}{\gamma_1}.$$

26. Sea γ una curva *cerrada* en el plano complejo (no necesariamente diferenciable) con un intervalo de parámetros $[0, 2\pi]$, tal que $\gamma(t) \neq 0$ para cada $t \in [0, 2\pi]$.

Elíjase $\delta > 0$ de tal forma que $|\gamma(t)| > \delta$ para todo $t \in [0, 2\pi]$. Si P_1 y P_2 son polinomios trigonométricos tales que $|P_j(t) - \gamma(t)| < \delta/4$ para todo $t \in [0, 2\pi]$ (su existencia la asegura el Teorema 8.15), demostrar que

$$\text{Índ } (P_1) = \text{Índ } (P_2)$$

aplicando el Ejercicio 25.

Definir este valor común como Índ (γ).

Demostrar que las proposiciones de los Ejercicios 24 y 25 se cumplen sin ninguna suposición de diferenciabilidad.

27. Sea f una función compleja continua definida en el plano complejo. Supóngase que hay un entero positivo n y un número complejo $c \neq 0$ tales que

$$\lim_{|z|\to\infty} z^{-n}f(z)=c.$$

Demostrar que $f(z) = 0$ para al menos un número complejo z.

Nótese que esto es una generalización del Teorema 8.8.

Sugerencia: Admítase que $f(z) \neq 0$ para todo z, y definir

$$\gamma_r(t)=f(re^{it})$$

si $0 \leq r < \infty$, $0 \leq t \leq 2\pi$ para demostrar que las siguientes proposiciones acerca de las curvas γ_r, son ciertas:

(*a*) Índ $(\gamma_0) = 0$.

(*b*) Índ $(\gamma_r) = n$ para todo r suficientemente grande.

(*c*) Índ (γ_r) es una función de r continua, sobre $[0, \infty)$.

[En (*b*) y (*c*), usar la última parte del Ejercicio 26.)

Mostrar que (*a*), (*b*) y (*c*) se contradicen, debido a que $n > 0$.

28. Sea $\bar{D}$ el disco unidad cerrado en el plano complejo. (Entonces $z \in \bar{D}$ si y solo si $|z| \leq 1$.) Sea g una aplicación continua de $\bar{D}$ en el círculo unidad T. (Entonces, $|g(z)| = 1$ para cada $z \in \bar{D}$.)

Demostrar que $g(z) = -z$ para al menos un $z \in T$.

Sugerencia: Para $0 \leq r \leq 1$, $0 \leq t \leq 2\pi$, hacer

$$\gamma_r(t)=g(re^{it}),$$

y poner $\psi(t) = e^{-it}\gamma_1(t)$. Si $g(z) \neq -z$ para cada $z \in T$, entonces $\psi(t) \neq -1$ para cada $t \in [0, 2\pi]$. De aquí que $(\psi) = 0$ por los Ejercicios 24 y 26. Se deduce que Ind $(\gamma_1) = 1$. Pero Ind $(\gamma_0) = 0$. Obtener una contradicción, como en el Ejercicio 27.

29. Demostrar que cada aplicación continua f de $\bar{D}$ en $\bar{D}$ tiene un punto fijo en $\bar{D}$.

Este es el teorema de punto fijo de Brouwer, para el caso de 2 dimensiones.)

Sugerencia: Suponer $f(z) \neq z$ para cada $z \in \bar{D}$. Asociar a cada $z \in \bar{D}$ el punto $g(z) \in T$ que se localiza sobre el rayo que parte de $f(z)$ y pasa por z. Entonces g mapea $\bar{D}$ en $T, g(z) = z$ si $z \in T$, y g es continua, porque

$$g(z) = z - s(z)[f(z) - z],$$

donde $s(z)$ es la única raíz no negativa de cierta ecuación cuadrática cuyos coeficientes son funciones continuas de f y z.

Aplicar el Ejercicio 28.

30. Mediante la fórmula de Stirling demostrar que

$$\lim_{x\to\infty} \frac{\Gamma(x + c)}{x^c\Gamma(x)} = 1$$

para cada constante real c.

31. Para la demostración del Teorema 7.26, comprobamos que

$$\int_{-1}^{1} (1 - x^2)^n \, dx \geq \frac{4}{3\sqrt{n}}$$

para $n = 1, 2, 3, \ldots$. Apoyándose en el Teorema 8.20 y Ejercicio 30, mostrar el siguiente resultado más preciso

$$\lim_{n\to\infty} \sqrt{n} \int_{-1}^{1} (1 - x^2)^n \, dx = \sqrt{\pi}.$$

9

FUNCIONES DE VARIAS VARIABLES

TRANSFORMACIONES LINEALES

Comenzamos este capítulo con un estudio de conjuntos de vectores en un n-espacio euclidiano R^n. Las propiedades algebraicas presentadas aquí, se extienden sin variación a espacios de vectores de dimensión finita sobre cualquier campo de escalares. Sin embargo, para nuestro objeto es totalmente suficiente permanecer en el marco habitual proporcionado por los espacios euclidianos.

9.1 Definiciones

(a) Un conjunto $X \subset R^n$ es un *espacio vectorial* si $\mathbf{x} + \mathbf{y} \in X$ y $c\mathbf{x} \in X$ para todo $\mathbf{x} \in X$, $\mathbf{y} \in X$, y para todo escalar c.

(b) Si $\mathbf{x}_1, \ldots, \mathbf{x}_k \in R^n$ y $c_1, \ldots, c_k$ son magnitudes escalares, al vector $c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k$ se le llama *combinación lineal de* $\mathbf{x}_1, \ldots, \mathbf{x}_k$. Si $S \subset R^n$ y si E es el conjunto de todas las combinaciones lineales de elementos de S, diremos que S *genera* a E, o que E es el *sistema de generadores* de S.

Se observará que todo sistema de generadores es un espacio vectorial.

(c) Se dice que un conjunto constituido por los vectores $\mathbf{x}_1, \ldots, \mathbf{x}_k$ (usaremos para tal conjunto la notación $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$) es *independiente*

(*linealmente independiente*), si la relación $c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k = 0$ implica que $c_1 = \cdots = c_k = 0$. En otro caso, se dice que $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ *es dependiente* (*linealmente dependiente*).

Obsérvese que ningún conjunto independiente contiene el vector nulo.

(*d*) Si un espacio vectorial X contiene un conjunto independiente de r vectores, pero no contiene un conjunto independiente de $r + 1$ vectores, diremos que tiene dimensión r y escribiremos: dim $X = r$.

El conjunto constituido por **0** solamente, es un espacio vectorial; su dimensión es **0**.

(*e*) De un subconjunto independiente de un espacio vectorial X que genera a X, se dice que es una *base* de X.

Obsérvese que si $B = \{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ es una base de X, todo $\mathbf{x} \in X$ tiene una representación única de la forma $\mathbf{x} = \Sigma c_j\mathbf{x}_j$. Tal representación existe, pues B genera a X y es única por ser B independiente. A los números $c_1, \ldots, c_r$ se le llama *coordenadas* de $\mathbf{x}$ con respecto a la base B.

El ejemplo más conocido de bases, es el conjunto $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, siendo $\mathbf{e}_j$ el vector en R^n cuya coordenada j-ésima es 1 y cuyas otras coordenadas son todas 0. Si $\mathbf{x} \in R^n$, $\mathbf{x} = (x_1, \ldots, x_n)$, entonces $\mathbf{x} = \Sigma x_j\mathbf{e}_j$. Llamaremos a $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la *base canónica, estándar o natural* de R^n.

9.2 Teorema *Sea r un entero positivo. Si un espacio vectorial X está engendrado por un conjunto r de vectores,* dim $X \leq r$.

Demostración Si no fuera cierto, existiría un espacio vectorial X que contendría un conjunto independiente $Q = \{\mathbf{y}_1, \ldots, \mathbf{y}_{r+1}\}$ que estaría engendrado por un conjunto S_0 de r vectores.

Supongamos que $0 \leq i \leq r$ y que se ha formado un conjunto S_i que genera a X y que consta de todos los $\mathbf{y}_j$ con $1 \leq j \leq i$ más una cierta colección de $r - i$ miembros de S_0, por ejemplo $\mathbf{x}_1, \ldots, \mathbf{x}_{r-i}$. (En otras palabras, S_i se obtiene de S_0 sustituyendo i de sus elementos por elementos de Q, sin modificar el sistema de generadores.) Como S_i genera a X, $\mathbf{y}_{i+1}$ pertenece al sistema de generadores de S_i; por tanto, existen escalares $a_1, \ldots, a_{i+1}, b_1, \ldots, b_{r-i}$, con $a_{i+1} = 1$, tales que

$$\sum_{j=1}^{i+1} a_j\mathbf{y}_j + \sum_{k=1}^{r-i} b_k\mathbf{x}_k = \mathbf{0}.$$

Si todo b_k fuese 0, la independencia de Q obligaría a todo a_j a ser 0, lo que es una contradicción. Se deduce que algún $\mathbf{x}_k \in S_i$ es una combinación lineal de los otros elementos de $T_i = S_i \cup \{\mathbf{y}_{i+1}\}$. Quitemos este $\mathbf{x}_k$ de T_i y llamemos S_{i+1} al conjunto que queda. Entonces, S_{i+1} ge-

nera el *mismo conjunto que* T_i, *es decir* X, *de modo que* S_{i+1} *tiene* las propiedades postuladas para S_i con $i + 1$ en lugar de i.

Partiendo de S_0, construimos así conjuntos $S_1, \ldots, S_r$. El último de éstos, consta de $\mathbf{y}_1, \ldots, \mathbf{y}_r$, y su construcción demuestra que genera a X. Pero Q es independiente; por tanto, $\mathbf{y}_{r+1}$ no pertenece al sistema de generadores de S_r. Esta contradicción demuestra el teorema.

Corolario dim $R^n = n$.

Demostración Como $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ genera R^n, el teorema demuestra que dim $R^n \leq n$. Como $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ es independiente, dim $R^n \geq n$.

9.3 Teorema *Supongamos que* X *es un espacio vectorial, y que* dim $X = n$.

(a) *Un conjunto* E *de* n *vectores en* X *genera a* X *si, y solo si* E *es independiente.*
(b) X *tiene una base, y toda base consta de* n *vectores.*
(c) *Si* $1 \leq r \leq n$ *y* $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$ *es un conjunto independiente en* X, *éste tiene una base que contiene a* $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$.

Demostración Supongamos que $E = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$. Como dim $X = n$, el conjunto $\{\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y}\}$ es dependiente, para todo $\mathbf{y} \in X$. Si E es independiente, se sigue que $\mathbf{y}$ pertenece al sistema de generadores de E, por lo que E genera a X. Inversamente, si E es dependiente, se puede suprimir uno de sus elementos sin cambiar el sistema de generadores de E. Por tanto, E no puede engendrar a X, por el Teorema 9.2, lo que demuestra (a).

Como dim $X = n$, X contiene un conjunto independiente de n vectores, y (a) muestra que dicho conjunto es una base de X; ahora, de 9.1(d) y 9.2 se deduce (b).

Para demostrar (c), sea $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ una base de X. El conjunto

$$S = \{\mathbf{y}_1, \ldots, \mathbf{y}_r, \mathbf{x}_1, \ldots, \mathbf{x}_n\}$$

genera X y es dependiente, pues contiene más de n vectores. El argumento utilizado en la demostración del Teorema 9.2 muestra que una de las $\mathbf{x}$ es una combinación lineal de los otros elementos de S. Si suprimimos este $\mathbf{x}$ de S, el conjunto que queda sigue engendrando a X. Este proceso puede repetirse r veces y conduce a una base de X que contiene a $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$, por ($a$).

9.4 Definiciones Se dice que una aplicación A de un espacio vectorial X en un espacio vectorial Y es una *transformación lineal* si

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2, \qquad A(c\mathbf{x}) = cA\mathbf{x}$$

para todo $\mathbf{x}$, $\mathbf{x}_1$, $\mathbf{x}_2 \in X$ y todo escalar c. Nótese que, a menudo, se escribe $A\mathbf{x}$ en lugar de $A(\mathbf{x})$, si A es lineal.

Se observará que $A\mathbf{0} = \mathbf{0}$, si A es lineal. Obsérvese también que una transformación lineal A de X en Y se determina completamente por medio de su acción sobre cualquier base: si $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ es una base de X, entonces cada $\mathbf{x} \in X$ tiene una representación única de la forma

$$\mathbf{x} = \sum_{i=1}^{n} c_i \mathbf{x}_i,$$

y la linealidad de A nos permite calcular $A\mathbf{x}$ a partir de los vectores $A\mathbf{x}_1, \ldots, A\mathbf{x}_n$ y las coordenadas $c_1, \ldots, c_n$ por medio de la fórmula

$$A\mathbf{x} = \sum_{i=1}^{n} c_i A\mathbf{x}_i.$$

A las transformaciones lineales de X en X se les llama frecuentemente *operadores lineales* en X. Si A es un operador lineal en X tal que: (i) es uno-a-uno, y (ii) aplica X sobre X, decimos que A es *invertible*. En este caso, podemos definir un operador A^{-1} en X con la condición de ser $A^{-1}(A\mathbf{x}) = \mathbf{x}$ para todo $\mathbf{x} \in X$. Es trivial comprobar que entonces se tiene también $A(A^{-1}\mathbf{x}) = \mathbf{x}$, para todo $\mathbf{x} \in X$ y que A^{-1} es lineal.

Una propiedad importante referente a los operadores lineales en espacios vectoriales de dimensión finita, es que cada una de las condiciones anteriores (i) e (ii) implica la otra.

9.5 Teorema *Un operador lineal A en un espacio vectorial X de dimensión finita, es uno-a-uno si y solo si el rango de de A es todo X.*

Demostración Sea $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ una base de X. La linealidad de A demuestra que su rango $\mathscr{R}(A)$ es el sistema de generadores del conjunto $Q = \{A\mathbf{x}_1, \ldots, A\mathbf{x}_n\}$. Por tanto, del Teorema 9.3(a) deducimos que $\mathscr{R}(A) = X$ si, y solo si Q es independiente. Tenemos que demostrar que esto sucede si, y solo si A es uno-a-uno.

Supongamos que A es uno-a-uno y $\Sigma c_i A\mathbf{x}_i = \mathbf{0}$. Entonces, $A(\Sigma c_i \mathbf{x}_i) = 0$, por lo que $\Sigma c_i \mathbf{x}_i = 0$ y, por tanto $c_1 = \cdots = c_n = 0$, llegando a la conclusión de que Q es independiente.

Inversamente, supongamos que Q es independiente y $A(\Sigma c_i \mathbf{x}_i) = 0$. Entonces, $\Sigma c_i A\mathbf{x}_i = 0$, por lo que $c_1 = \cdots = c_n = 0$ y deducimos que $A\mathbf{x} = 0$ solo si $\mathbf{x} = 0$. Si, ahora, $A\mathbf{x} = A\mathbf{y}$ será $A(\mathbf{x} - \mathbf{y}) = A\mathbf{x} - A\mathbf{y} = 0$, de forma que $\mathbf{x} - \mathbf{y} = 0$, lo que nos dice que A es uno-a-uno.

9.6 Definiciones

(*a*) Sea $L(X, Y)$ el conjunto de todas las transformaciones lineales del espacio vectorial X en el espacio vectorial Y. En lugar de $L(X,X)$ escribiremos, simplemente, $L(X)$. Si A_1, $A_2 \in L(Y,X)$ y c_1, c_2 son escalares, definiremos $c_1A_1 + c_2A_2$ por

$$(c_1A_1 + c_2A_2)\mathbf{x} = c_1A_1\mathbf{x} + c_2A_2\mathbf{x} \qquad (\mathbf{x} \in X).$$

Es claro, entonces, que $c_1A_1 + c_2A_2 \in L(X, Y)$.
(*b*) Si X, Y, Z son espacios vectoriales, si $A \in L(X,Y)$ y $B \in L(Y,Z)$ definiremos su *producto* BA como la *composición* de A y B:

$$(BA)\mathbf{x} = B(A\mathbf{x}) \qquad (\mathbf{x} \in X).$$

Entonces, $BA \in L(X,Z)$.

Obsérvese que BA no es necesariamente igual a AB, incluso cuando $X = Y = Z$.
(*c*) Para $A \in L(R^n, R^m)$, definiremos la *norma* $\|A\|$ de A como el sup de todos los números $|A\mathbf{x}|$, donde $\mathbf{x}$ tiene como rango todos los vectores R^n con $|\mathbf{x}| \leq 1$.

$$|A\mathbf{x}| \leq \|A\| \, |\mathbf{x}|$$

se cumple para todo $\mathbf{x} \in R^n$. Además, si λ es tal que $|A\mathbf{x}| \leq \lambda|\mathbf{x}|$ para todo $\mathbf{x} \in R^n$, entonces $\|A\| \leq \lambda$.

9.7 Teorema

(*a*) *Si $A \in L(R^n, R^m)$, entonces $\|A\| < \infty$ y A es una aplicación uniformemente continua de R^n en R^m.*
(*b*) *Si A, $B \in L(R^n, R^m)$ y c es un escalar, será*

$$\|A + B\| \leq \|A\| + \|B\|, \qquad \|cA\| = |c| \, \|A\|.$$

Con la distancia entre A y B definida por $\|A - B\|$, $L(R^n, R^m)$ es un espacio métrico.
(*c*) *Si $A \in L(R^n, R^m)$ y $B \in L(R^m, R^k)$, entonces*

$$\|BA\| \leq \|B\| \, \|A\|.$$

Demostración

(*a*) Sea $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base natural en R^n y supongamos que $\mathbf{x} = \Sigma c_i\mathbf{e}_i$, $|\mathbf{x}| \leq 1$, de modo que $|c_i| \leq 1$ para $i = 1, \ldots, n$. Será:

$$|A\mathbf{x}| = \left|\sum c_i A\mathbf{e}_i\right| \leq \sum |c_i|\ |A\mathbf{e}_i| \leq \sum |A\mathbf{e}_i|$$

de forma que

$$\|A\| \leq \sum_{i=1}^{n} |A\mathbf{e}_i| < \infty.$$

Como $|A\mathbf{x} - A\mathbf{y}| \leq \|A\|\ |\mathbf{x} - \mathbf{y}|$ si $\mathbf{x}, \mathbf{y} \in R^n$, vemos que A es uniformemente continua.
(b) La desigualdad de (b) se deduce de

$$|(A + B)\mathbf{x}| = |A\mathbf{x} + B\mathbf{x}| \leq |A\mathbf{x}| + |B\mathbf{x}| \leq (\|A\| + \|B\|)\,|\mathbf{x}|.$$

Del mismo modo se demuestra la segunda parte de (b). Si

$$A,\ B,\ C \in L(R^n, R^m),$$

tenemos la desigualdad del triángulo

$$\|A - C\| = \|(A - B) + (B - C)\| \leq \|A - B\| + \|B - C\|,$$

y se comprueba fácilmente que $\|A - B\|$ tiene las restantes propiedades de una métrica. (Definición 2.15).
(c) *Finalmente,* (c) se deduce de

$$|(BA)\mathbf{x}| = |B(A\mathbf{x})| \leq \|B\|\ |A\mathbf{x}| \leq \|B\|\ \|A\|\ |\mathbf{x}|.$$

Como ahora tenemos métricas en los espacios $L(R^n, R^m)$, los conceptos de conjunto abierto, continuidad, etc. Tienen sentido para estos espacios. El próximo teorema utiliza estos conceptos.

9.8 Teorema *Sea Ω el conjunto de todos los operadores lineales invertibles en R^n.*

(a) *Si $A \in \Omega$, $B \in L(R^n)$, y*

$$\|B - A\| \cdot \|A^{-1}\| < 1,$$

entonces $B \in \Omega$.
(b) *Ω es un subconjunto abierto de $L(R^n)$, y la aplicación $A \to A^{-1}$ es continua en Ω.*

(Además, es evidente, una aplicación 1-1 de Ω sobre sí mismo, que es su propio inverso.)

Demostración

(a) Poniendo $\|A^{-1}\| = 1/\alpha$, y $\|B - A\| = \beta$. Entonces $\beta < \alpha$. Para cada $\mathbf{x} \in R^n$,

$$\begin{aligned}\alpha|\mathbf{x}| &= \alpha|A^{-1}A\mathbf{x}| \le \alpha\|A^{-1}\| \cdot |A\mathbf{x}| \\ &= |A\mathbf{x}| \le |(A - B)\mathbf{x}| + |B\mathbf{x}| \le \beta|\mathbf{x}| + |B\mathbf{x}|,\end{aligned}$$

así que

(1) $$(\alpha - \beta)|\mathbf{x}| \le |B\mathbf{x}| \qquad (\mathbf{x} \in R^n).$$

Como $\alpha - \beta > 0$, (1) muestra que $B\mathbf{x} \neq 0$ si $\mathbf{x} \neq 0$. De esto se tiene que B es 1-1. Del Teorema 9.5, $B \in \Omega$. Esto se cumple para todo B con $\|B - A\| < \alpha$. Entonces se tiene (a) y el hecho de que Ω es abierto.

(b) Ahora, se reemplaza $\mathbf{x}$ por $B^{-1}\mathbf{y}$ en (1). La desigualdad resultante

(2) $$(\alpha - \beta)|B^{-1}\mathbf{y}| \le |BB^{-1}\mathbf{y}| = |\mathbf{y}| \qquad (\mathbf{y} \in R^n)$$

muestra que $\|B^{-1}\| \le (\alpha - \beta)^{-1}$. La identidad

$$B^{-1} - A^{-1} = B^{-1}(A - B)A^{-1},$$

combinada con el Teorema 9.7(c), implica por tanto que

$$\|B^{-1} - A^{-1}\| \le \|B^{-1}\| \; \|A - B\| \; \|A^{-1}\| \le \frac{\beta}{\alpha(\alpha - \beta)}.$$

Esto establece la afirmación de continuidad que se hizo en (b), debido a que $\beta \to 0$ cuando $B \to A$.

9.9 Matrices Supongamos que $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ y $\{\mathbf{y}_1, \ldots, \mathbf{y}_m\}$ son bases de espacios vectoriales X, Y, respectivamente. Cada $A \in L(X, Y)$ determinará un conjunto de números a_{ij} tales que

(3) $$A\mathbf{x}_j = \sum_{i=1}^{m} a_{ij}\mathbf{y}_i \qquad (1 \le j \le n).$$

Es conveniente colocar estos números en un cuadro de m filas y n columnas, llamado matriz m por n.

$$[A] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Se observará que las coordenadas a_{ij} del vector $A\mathbf{x}_j$ (con respecto a la base $\{\mathbf{y}_1, \ldots, \mathbf{y}_m\}$ aparecen en la columna j-ésima de $[A]$. A los vectores $A\mathbf{x}_j$ se les llama a veces *vectores columna* de $[A]$. Con esta terminología, *el rango de A está engendrado por los vectores columna de* $[A]$.

Si $\mathbf{x} = \Sigma c_j \mathbf{x}_j$, la linealidad de A, combinada con (3) muestra que

(4)
$$A\mathbf{x} = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} c_j \right) \mathbf{y}_i .$$

Así pues, las coordenadas de $A\mathbf{x}$ son $\Sigma \; a_{ij} c_j$. Nótese que en (3) la sumación se extiende sobre el primer subíndice de a_{ij}; cuando calculamos las coordenadas, sumamos respecto al segundo.

Supongamos ahora, que se da una matriz m por n, con términos reales a_{ij}. Si, entonces, A está definido por (4), es evidente que $A \in L(X, Y)$ y que $[A]$ es la matriz dada. Así, hay una correspondencia natural 1-1 entre $L(X, Y)$ y el conjunto de todas las matrices reales m por n. Sin embargo, recalcamos que $[A]$ depende no solo de A sino también de la elección de bases en x y Y. El mismo A puede dar origen a muchas matrices diferentes si cambiamos las bases, y viceversa. No llevaremos más adelante esta observación, pues habitualmente trabajaremos con bases fijas. (En 9.37 se pueden encontrar algunas observaciones sobre este tema.)

Si Z es un tercer espacio vectorial, con base $\{\mathbf{z}_1, \ldots, \mathbf{z}_p\}$, si A está dada por (3) y si

$$B\mathbf{y}_i = \sum_k b_{ki} \mathbf{z}_k, \qquad (BA)\mathbf{x}_j = \sum_k c_{kj} \mathbf{z}_k,$$

entonces $A \in L(X, Y)$, $B \in L(Y, Z)$, $BA \in L(X, Z)$, y como

$$\begin{aligned} B(A\mathbf{x}_j) &= B \sum_i a_{ij} \mathbf{y}_i = \sum_i a_{ij} B\mathbf{y}_i \\ &= \sum a_{ij} \sum_k b_{ki} \mathbf{z}_k = \sum_k \left(\sum_i b_{ki} a_{ij} \right) \mathbf{z}_k , \end{aligned}$$

la independencia de $\{\mathbf{z}_1, \ldots, \mathbf{z}_p\}$ implica que

(5)
$$c_{kj} = \sum_i b_{ki} a_{ij} \qquad (1 \leq k \leq p, \; 1 \leq j \leq n).$$

Esto muestra el modo de calcular la matriz p por n $[BA]$ a partir de $[B]$ y $[A]$. Si definimos el producto $[B][A]$ como $[BA]$, entonces (5) describe la regla normal de multiplicación de matrices.

Finalmente, supongamos que $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ y $\{\mathbf{y}_1, \ldots, \mathbf{y}_m\}$ son bases estándar de R^n y R^m, y A está dado por (4). La desigualdad de Schwarz muestra que

$$|A\mathbf{x}|^2 = \sum_i \left(\sum_j a_{ij} c_j \right)^2 \leq \sum_i \left(\sum_j a_{ij}^2 \cdot \sum_j c_j^2 \right) = \sum_{i,j} a_{ij}^2 |\mathbf{x}|^2 .$$

Así, pues,

$$\|A\| \leq \left\{\sum_{i,j} a_{ij}^2\right\}^{1/2}. \tag{6}$$

Si aplicamos (6) a $B - A$ en lugar de A, donde $A, B \in L(R^n, R^m)$, vemos que si los elementos de la matriz a_{ij} son funciones continuas de un parámetro, esto mismo es cierto para A. Más preciso:

Si S es un espacio métrico, si $a_{11}, \ldots, a_{mn}$ son funciones continuas reales en S, y si para cada $p \in S$, A_p es la transformación lineal de R^n en R^m, cuya matriz tiene términos $a_{ij}(p)$, entonces la aplicación $p \to A_p$ es una aplicación continua de S en $L(R^n, R^m)$.

DIFERENCIACIÓN

9.10 Preliminares Con el fin de obtener una definición de la derivada de una función cuyo dominio es R^n (o un subconjunto abierto de R^n), volvamos al caso conocido $n = 1$, y veamos cómo interpretar la derivada en este caso, de tal manera que el resultado pueda ampliarse de manera natural cuando $n > 1$.

Si f es una función real con dominio $(a,b) \subset R^1$ y si $x \in (a,b)$, entonces $f'(x)$ se define comúnmente como el número real

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}, \tag{7}$$

por supuesto, siempre que este límite exista. Por lo que

$$f(x+h) - f(x) = f'(x)h + r(h) \tag{8}$$

donde el "residuo" $r(h)$ es pequeño, entendiéndose por esto que

$$\lim_{h \to 0} \frac{r(h)}{h} = 0. \tag{9}$$

Nótese que (8) expresa la diferencia $f(x + h) - f(x)$ como la suma de la *función lineal* que lleva h a $f'(x)h$, más un pequeño residuo.

Puede considerarse por tanto, que la derivada de f en x no es un número real, sino un operador lineal sobre R^1 que lleva a h a $f'(x)h$.

[Obsérvese que cada número real α da lugar a un operador lineal sobre R^1; el operador en cuestión es sencillamente la multiplicación por α. De manera inversa, toda función lineal que lleva R^1 en R^1 es la multiplicación por algún número real. Esta correspondencia 1-1 natural entre R^1 y $L(R^1)$ es la que motiva las proposiciones anteriores.]

Considérese ahora una función $\mathbf{f}$ que mapea $(a,b) \subset R^1$ en R^m. En este caso, $\mathbf{f}'(x)$ se definió como el vector $\mathbf{y} \in R^m$ (si existe uno) para el cual

$$\lim_{h\to 0} \left\{\frac{\mathbf{f}(x+h)-\mathbf{f}(x)}{h} - \mathbf{y}\right\} = \mathbf{0}. \tag{10}$$

Esto puede volverse a escribir en la forma

$$\mathbf{f}(x+h) - \mathbf{f}(x) = h\mathbf{y} + \mathbf{r}(h), \tag{11}$$

en donde $\mathbf{r}(h)/h \to \mathbf{0}$ cuando $h \to 0$. El término principal en el miembro derecho de (11) es, otra vez, una función *lineal* de h. Cada $\mathbf{y} \in R^n$ induce una transformación lineal de R^1 en R^m, asociando a cada $h \in R^1$ el vector $h\mathbf{y} \in R^m$. Esta identificación de R^m con $L(R^1, R^m)$ permite considerar a $\mathbf{f}'(x)$ como un miembro de $L(R^1, R^m)$.

Entonces, si $\mathbf{f}$ es un mapeo diferenciable de $(a,b) \subset R^1$ en R^m, y si $x \in (a,b)$, entonces $\mathbf{f}(x)$ es una transformación lineal de R^1 en R^m que satisface

$$\lim_{h\to 0} \frac{\mathbf{f}(x+h)-\mathbf{f}(x)-\mathbf{f}'(x)h}{h} = \mathbf{0}, \tag{12}$$

o de manera equivalente,

$$\lim_{h\to 0} \frac{|\mathbf{f}(x+h)-\mathbf{f}(x)-\mathbf{f}'(x)h|}{|h|} = 0. \tag{13}$$

Podemos ahora considerar el caso $n > 1$.

9.11 Definición Supongamos que E es un conjunto abierto en R^n, que $\mathbf{f}$ mapea E en R^m, y $\mathbf{x} \in E$. Si existe una transformación lineal A de R^n en R^m, tal que

$$\lim_{\mathbf{h}\to 0} \frac{|\mathbf{f}(\mathbf{x}+\mathbf{h})-\mathbf{f}(\mathbf{x})-A\mathbf{h}|}{|\mathbf{h}|} = 0, \tag{14}$$

decimos que $\mathbf{f}$ es *diferenciable* en $\mathbf{x}$, y escribimos

$$\mathbf{f}'(\mathbf{x}) = A. \tag{15}$$

Si $\mathbf{f}$ es diferenciable en todo $\mathbf{x} \in E$ decimos que es *diferenciable* en E.

En (14) se sobreentiende, naturalmente, que $\mathbf{h} \in R^n$; si $|\mathbf{h}|$ es suficientemente pequeño, entonces $\mathbf{x} + \mathbf{h} \in E$, pues, E es abierto. Así, pues, $\mathbf{f}(\mathbf{x} + \mathbf{h})$ está definida, $\mathbf{f}(\mathbf{x} + \mathbf{h}) \in R^m$, y como $A \in L(R^n, R^m)$, $A\mathbf{h} \in R^m$. Así,

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - A\mathbf{h} \in R^m.$$

La norma del numerador de (5) es la de R^m; en el denominador, tenemos la norma $-R^n$ de $\mathbf{h}$.

Hay un problema obvio de unicidad que tiene que establecerse antes de seguir adelante.

9.12 Teorema *Supongamos que E y $\mathbf{f}$ son como en la Definición 9.11,* $\mathbf{x} \in E$, *y se cumple* (14) *con* $A = A_1$ *y con* $A = A_2$, *Entonces,* $A_1 = A_2$.

Demostración Si $B = A_1 - A_2$, la desigualdad

$$|B\mathbf{h}| \leq |\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - A_1\mathbf{h}| + |\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - A_2\mathbf{h}|$$

muestra que $|B\mathbf{h}|/|\mathbf{h}| \to 0$ cuando $\mathbf{h} \to 0$. Para $\mathbf{h} \neq 0$ fijado, se deduce que

(16) $$\frac{|B(t\mathbf{h})|}{|t\mathbf{h}|} \to 0 \quad \text{cuando} \quad t \to 0.$$

La linealidad de B demuestra que el primer miembro de (10) es independiente de t. Así, pues, $B\mathbf{h} = \mathbf{0}$ para todo $\mathbf{h} \in R^n$, de donde se deduce que $B = 0$.

9.13 Observaciones

(*a*) La relación (14) puede escribirse nuevamente en la forma

(17) $$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\mathbf{h} + \mathbf{r}(\mathbf{h})$$

donde el residuo $\mathbf{r}(\mathbf{h})$ es pequeño, en el sentido que

(18) $$\lim_{\mathbf{h}\to 0} \frac{|\mathbf{r}(\mathbf{h})|}{|\mathbf{h}|} = 0.$$

Podemos interpretar (17) diciendo como en la sección 9.10, que para $\mathbf{x}$ fijado y $\mathbf{h}$ pequeño,

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})$$

es, aproximadamente, igual a $\mathbf{f}'(\mathbf{x})\mathbf{h}$, esto es, al valor de una función lineal aplicada a $\mathbf{h}$.

(*b*) Supongamos que $\mathbf{f}$ y E son como en la Definición 9.11, y que $\mathbf{f}$ es diferenciable en E. Para cada $\mathbf{x} \in E$, $\mathbf{f}'(\mathbf{x})$ es una función, esto es, una transformación lineal de R^n en R^m. Pero $\mathbf{f}'$ es también una función: aplica E en $L(R^n, R^m)$.

(*c*) Una ojeada a (17) muestra que $\mathbf{f}$ es continua en todo punto, en el cual es diferenciable.
(*d*) La derivada definida por (14) o por (17) se llama frecuentemente *derivada total* de $\mathbf{f}$ en $\mathbf{x}$ o *diferencial* de $\mathbf{f}$ en $\mathbf{x}$, para distinguirlas de las derivadas parciales que posteriormente se presentarán.

9.14 Ejemplo Se han definido las derivadas de funciones que llevan R^n a R^m como transformaciones lineales de R^n en R^m. ¿Qué es la derivada de dicha transformación lineal? La respuesta es muy sencilla.

Si $A \in L(R^n, R^m)$ *y si* $\mathbf{x} \in R^n$, *entonces*

$$A'(\mathbf{x}) = A. \tag{19}$$

Nótese que $\mathbf{x}$ aparece en el miembro izquierdo de (19), pero no en el derecho. Los dos lados de (19) son miembros de $L(R^n, R^m)$, puesto que $A\mathbf{x} \in R^m$.

La demostración de (19) es trivial, ya que

$$A(\mathbf{x} + \mathbf{h}) - A\mathbf{x} = A\mathbf{h}, \tag{20}$$

por la linealidad de A. Haciendo $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$, el numerador en (14) es por esto 0, para cada $\mathbf{h} \in R^n$. En (17) $\mathbf{r}(\mathbf{h}) = \mathbf{0}$.

Ahora se ampliará la regla de la cadena (Teorema 5.5), para la situación presente.

9.15 Teorema *Supongamos que E es un conjunto abierto en R^n,* $\mathbf{f}$ *mapea E en R^m,* $\mathbf{f}$ *es diferenciable en* $\mathbf{x}_0 \in E$, $\mathbf{g}$ *mapea un conjunto abierto que contiene a* $\mathbf{f}(E)$ *en R^k, y* $\mathbf{g}$ *es diferenciable en* $\mathbf{f}(\mathbf{x}_0)$. *El mapeo* $\mathbf{F}$ *de E en R^k definido por*

$$\mathbf{F}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$$

es diferenciable en $\mathbf{x}_0$, *y*

$$\mathbf{F}'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0))\mathbf{f}'(\mathbf{x}_0). \tag{21}$$

En el segundo miembro de (21), tenemos el producto de dos transformaciones lineales, definido en 9.6.

Demostración Pongamos $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$; $A = \mathbf{f}'(\mathbf{x}_0)$; $B = \mathbf{g}'(\mathbf{y}_0)$, y definamos

$$\mathbf{u}(\mathbf{h}) = \mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) - A\mathbf{h},$$
$$\mathbf{v}(\mathbf{k}) = \mathbf{g}(\mathbf{y}_0 + \mathbf{k}) - \mathbf{g}(\mathbf{y}_0) - B\mathbf{k},$$

para todo $\mathbf{h} \in R^n$ y $\mathbf{k} \in R^m$ para los cuales $\mathbf{f}(\mathbf{x}_0 + \mathbf{h})$ y $\mathbf{g}(\mathbf{y}_0 + \mathbf{k})$ están definidas. Entonces

(22) $$|\mathbf{u}(\mathbf{h})| = \varepsilon(\mathbf{h})|\mathbf{h}|, \qquad |\mathbf{v}(\mathbf{k})| = \eta(\mathbf{k})|\mathbf{k}|,$$

donde $\varepsilon(\mathbf{h}) \to 0$ cuando $\mathbf{h} \to \mathbf{0}$ y $\eta(\mathbf{k}) \to 0$ cuando $\mathbf{k} \to \mathbf{0}$.

Dado $\mathbf{h}$, se pone $\mathbf{k} = \mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0)$. Entonces

(23) $$|\mathbf{k}| = |A\mathbf{h} + \mathbf{u}(\mathbf{h})| \leq [\|A\| + \varepsilon(\mathbf{h})]\,|\mathbf{h}|,$$

y

$$\begin{aligned}\mathbf{F}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{F}(\mathbf{x}_0) - BA\mathbf{h} &= \mathbf{g}(\mathbf{y}_0 + \mathbf{k}) - \mathbf{g}(\mathbf{y}_0) - BA\mathbf{h} \\ &= B(\mathbf{k} - A\mathbf{h}) + \mathbf{v}(\mathbf{k}) \\ &= B\mathbf{u}(\mathbf{h}) + \mathbf{v}(\mathbf{k}).\end{aligned}$$

En consecuencia (22) y (23) implican, para $\mathbf{h} \neq \mathbf{0}$, que

$$\frac{|\mathbf{F}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{F}(\mathbf{x}_0) - BA\mathbf{h}|}{|\mathbf{h}|} \leq \|B\|\,\varepsilon(\mathbf{h}) + [\|A\| + \varepsilon(\mathbf{h})]\eta(\mathbf{k}).$$

Sea $\mathbf{h} \to \mathbf{0}$. Entonces $\varepsilon(\mathbf{h}) \to 0$. También $\mathbf{k} \to \mathbf{0}$, por (23), así que $\eta(\mathbf{k}) \to 0$. Se deduce que $\mathbf{F}(\mathbf{x}_0) = BA$, que es lo que asegura (21).

9.16 Derivadas parciales Considérese otra vez una función $\mathbf{f}$ que mapea un conjunto abierto $E \subset R^n$ en R^m. Sean $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ y $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ las bases estándar de R^n y R^m. Las *componentes de* $\mathbf{f}$ son las funciones reales $f_1, \ldots, f_m$ definidas como

(24) $$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{m} f_i(\mathbf{x})\mathbf{u}_i \qquad (\mathbf{x} \in E),$$

o de manera equivalente, por $f_i(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \cdot \mathbf{u}_i$, $1 \leq i \leq m$.

Para $\mathbf{x} \in E$, $1 \leq i \leq m$, $1 \leq j \leq n$, se define

(25) $$(D_j f_i)(\mathbf{x}) = \lim_{t \to 0} \frac{f_i(\mathbf{x} + t\mathbf{e}_j) - f_i(\mathbf{x})}{t},$$

siempre que el límite exista. Escribiendo $f_i(x_1, \ldots, x_n)$ en lugar de $f_i(\mathbf{x})$, se ve que $D_j f_i$ es la derivada de f_i con respecto a x_j, conservando las otras variables fijas. La notación

(26) $$\frac{\partial f_i}{\partial x_j}$$

se usa con frecuencia en lugar de $D_j f_i$, y a $D_j f_i$ se le llama una *derivada parcial.*

En muchos casos, cuando se trata con funciones de una variable en donde la existencia de la derivada es suficiente, la continuidad, o al menos, la acotabilidad de las derivadas parciales es necesaria para funciones de varias variables. Por ejemplo las funciones f y g que se describieron en el Ejercicio 7 del capítulo 4, no son continuas, aunque sus derivadas parciales existen en cada punto de R^2. Incluso para funciones continuas, la existencia de todas las derivadas parciales no implica la diferenciabilidad según la Definición 9.11; véanse los Ejercicios 6 y 14, y el Teorema 9.21.

No obstante, si se sabe que $\mathbf{f}$ es diferenciable en un punto $\mathbf{x}$, entonces sus derivadas parciales existen en $\mathbf{x}$ y determinan la transformación lineal $\mathbf{f}'(\mathbf{x})$ completamente:

9.17 Teorema *Supóngase que* $\mathbf{f}$ *mapea un conjunto abierto* $E \subset R$ *en* R^m, *y que* $\mathbf{f}$ *es diferenciable en un punto* $\mathbf{x} \in E$. *Entonces las derivadas parciales* $(D_j f_i)(\mathbf{x})$ *existen, y*

(27) $$\mathbf{f}'(\mathbf{x})\mathbf{e}_j = \sum_{i=1}^{m} (D_j f_i)(\mathbf{x})\mathbf{u}_i \qquad (1 \le j \le n).$$

Aquí, como en la sección 9.16, $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ y $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ son las bases estándar de R^n y R^m.

Demostración Con j fijo, y debido a que $\mathbf{f}$ es diferenciable en $\mathbf{x}$,

$$\mathbf{f}(\mathbf{x} + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}) = \mathbf{f}'(\mathbf{x})(t\mathbf{e}_j) + \mathbf{r}(t\mathbf{e}_j)$$

en donde $|\mathbf{r}(t\mathbf{e}_j)|/t \to 0$ cuando $t \to 0$. Por tanto, la linealidad de $\mathbf{f}(\mathbf{x})$ muestra que

(28) $$\lim_{t\to 0} \frac{\mathbf{f}(\mathbf{x} + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x})}{t} = \mathbf{f}'(\mathbf{x})\mathbf{e}_j.$$

Si ahora se representa $\mathbf{f}$ en términos de sus componentes, como en (24), entonces (28) se convierte en

(29 $$\lim_{t\to 0} \sum_{i=1}^{m} \frac{f_i(\mathbf{x} + t\mathbf{e}_j) - f_i(\mathbf{x})}{t} \mathbf{u}_i = \mathbf{f}'(\mathbf{x})\mathbf{e}_j.$$

Se deduce que cada cociente tiene un límite en su suma cuando $t \to 0$ (véase Teorema 4.10), así que cada $(D_j f_i)(\mathbf{x})$ existe, y entonces (27) se deduce de (29).

He aquí algunas consecuencias del Teorema 9.17:

Sea $[\mathbf{f}'(\mathbf{x})]$ una matriz que representa $\mathbf{f}'(\mathbf{x})$ con respecto a las bases estándar, como se hizo en la sección 9.9.

Entonces $\mathbf{f}(\mathbf{x})\mathbf{e}_j$ es el j-ésimo vector columna de $[\mathbf{f}'(\mathbf{x})]$, y (27) muestra, por lo tanto, que el número $(D_j f_i)(\mathbf{x})$ ocupa el lugar en el i-ésimo renglón y la j-ésima columna de $[\mathbf{f}'(\mathbf{x})]$. Así

$$[\mathbf{f}'(\mathbf{x})] = \begin{bmatrix} (D_1 f_1)(\mathbf{x}) & \cdots & (D_n f_1)(\mathbf{x}) \\ \cdots & \cdots & \cdots \\ (D_1 f_m)(\mathbf{x}) & \cdots & (D_n f_m)(\mathbf{x}) \end{bmatrix}.$$

Si $\mathbf{h} = \Sigma h_j \mathbf{e}_j$ es cualquier vector en R^n, entonces (27) implica que

$$\mathbf{f}'(\mathbf{x})\mathbf{h} = \sum_{i=1}^{m} \left\{ \sum_{j=1}^{n} (D_j f_i)(\mathbf{x}) h_j \right\} \mathbf{u}_i. \tag{30}$$

9.18 Ejemplo Sea γ un mapeo diferenciable del segmento $(a,b) \subset R^1$ en un conjunto abierto $E \subset R^n$; en otras palabras, γ es una curva diferenciable en E. Sea f una función diferenciable de valores reales con dominio E. Entonces f es un mapeo diferenciable de E en R^1. Se define

$$g(t) = f(\gamma(t)) \quad (a < t < b). \tag{31}$$

La regla de la cadena asegura entonces que

$$g'(t) = f'(\gamma(t))\gamma'(t) \qquad (a < t < b). \tag{32}$$

Como $\gamma'(t) \in L(R^1, R^n)$ y $f'(\gamma(t)) \in L(R^n, R^1)$, (32) define a $g'(t)$ como un operador lineal sobre R^1. Esto coincide con el hecho de que g mapea (a,b) en R^1.

Sin embargo, $g'(t)$ también puede considerarse como un número real. (Esto se trató en la Sec. 9.10.) Este número se puede calcular en términos de las derivadas parciales de f y las derivadas de las componentes de γ, como se verá a continuación.

Con respecto a la base estándar $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ de R^n, $[\gamma(t)]$ es la matriz n por 1 (una "matriz columna") que tiene a $\gamma'_i(t)$ en el i-ésimo renglón, y donde $\gamma_1, \ldots, \gamma_n$ son las componentes de γ. Para cada $\mathbf{x} \in E$, $[f'(\mathbf{x})]$ es la matriz 1 por n (una "matriz renglón") que tiene a $(D_j f)(\mathbf{x})$ en la j-ésima columna. Por consiguiente, $[g'(t)]$ es la matriz 1 por 1 cuyo único elemento es el número real

$$g'(t) = \sum_{i=1} (D_i f)(\gamma(t))\gamma'_i(t). \tag{33}$$

Este es un caso especial de la regla de la cadena que se encuentra con mucha frecuencia. Se puede volver a escribir de la manera siguiente:

Con cada $\mathbf{x} \in E$ se asocia un vector, llamado "gradiente" de f en $\mathbf{x}$, definido como

(34) $$(\nabla f)(\mathbf{x}) = \sum_{i=1}^{n} (D_i f)(\mathbf{x})\mathbf{e}_i.$$

Como

(35) $$\gamma'(t) = \sum_{i=1}^{n} \gamma_i'(t)\mathbf{e}_i,$$

(33) puede escribirse en la forma

(36) $$g'(t) = (\nabla f)(\gamma(t)) \cdot \gamma'(t),$$

que es el producto escalar de los vectores $(\nabla f)(\gamma(t))$ y $\gamma'(t)$.

Fijemos ahora un $\mathbf{x} \in E$, y sea $\mathbf{u} \in R^n$ un vector unitario (es decir, $|\mathbf{u}| = 1$, y en especial γ de manera que

(37) $$\gamma(t) = \mathbf{x} + t\mathbf{u} \qquad (-\infty < t < \infty).$$

Entonces $\gamma'(t) = \mathbf{u}$ para cada t. Por consiguiente (36) muestra que

(38) $$g'(0) = (\nabla f)(\mathbf{x}) \cdot \mathbf{u}.$$

Por otro lado, (37) muestra que

$$g(t) - g(0) = f(\mathbf{x} + t\mathbf{u}) - f(\mathbf{x}).$$

De aquí que (38) produzca

(39) $$\lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{u}) - f(\mathbf{x})}{t} = (\nabla f)(\mathbf{x}) \cdot \mathbf{u}.$$

El límite en (39) se denomina comúnmente la *derivada direccional* de f en $\mathbf{x}$, en la dirección del vector unitario $\mathbf{u}$, y puede representarse por $(D_{\mathbf{u}} f)(\mathbf{x})$.

Si f y $\mathbf{x}$ son fijos, pero $\mathbf{u}$ varía, entonces (39) muestra que $D_{\mathbf{u}} f)(\mathbf{x})$ alcanza su máximo cuando $\mathbf{u}$ es un múltiplo escalar positivo de $(\nabla f)(\mathbf{x})$. (El caso $(\nabla f)(\mathbf{x}) = \mathbf{0}$ no se incluirá aquí.)

Si $\mathbf{u} = \Sigma u_i \mathbf{e}_i$, entonces (39) muestra que $(D_{\mathbf{u}} f)(\mathbf{x})$ puede expresarse en términos de las derivadas parciales de f en $\mathbf{x}$ por medio de la fórmula

(40) $$(D_{\mathbf{u}} f)(\mathbf{x}) = \sum_{i=1}^{n} (D_i f)(\mathbf{x}) u_i.$$

Algunas de estas ideas juegan un papel en el teorema siguiente.

9.19 Teorema *Supóngase que* **f** *mapea un conjunto abierto convexo* $E \subset R^n$ *en* R^m, **f** *es diferenciable en* E, *y hay un número real* M *tal que*

$$\|\mathbf{f}'(\mathbf{x})\| \leq M$$

para cada $\mathbf{x} \in E$. *Entonces*

$$|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})| \leq M|\mathbf{b} - \mathbf{a}|$$

para todo $\mathbf{a} \in E$, $\mathbf{b} \in E$.

Demostración Con $\mathbf{a} \in E$, $\mathbf{b} \in E$ fijos, se define

$$\gamma(t) = (1 - t)\mathbf{a} + t\mathbf{b}$$

Para todo $t \in R^1$ tal que $\gamma(t) \in E$. Como E es convexo, $\gamma(t) \in E$ si $0 \leq t \leq 1$. Poniendo

$$\mathbf{g}(t) = \mathbf{f}(\gamma(t)).$$

Entonces

$$\mathbf{g}'(t) = \mathbf{f}'(\gamma(t))\gamma'(t) = \mathbf{f}'(\gamma(t))(\mathbf{b} - \mathbf{a}),$$

así que

$$|\mathbf{g}'(t)| \leq \|\mathbf{f}'(\gamma(t))\| \, |\mathbf{b} - \mathbf{a}| \leq M|\mathbf{b} - \mathbf{a}|$$

para todo $t \in [0, 1]$. Por el Teorema 5.19,

$$|\mathbf{g}(1) - \mathbf{g}(0)| \leq M|\mathbf{b} - \mathbf{a}|.$$

Pero $\mathbf{g}(0) = \mathbf{f}(a)$ y $\mathbf{g}(1) = \mathbf{f}(b)$. Esto completa la demostración.

Corolario *Si además,* $\mathbf{f}'(\mathbf{x}) = \mathbf{0}$ *para todo* $\mathbf{x} \in E$, *entonces* **f** *es constante.*

Demostración Para demostrar esto, nótese que las hipótesis del teorema se cumplen ahora con $M = 0$.

9.20 Definición Un mapeo diferenciable **f** de un conjunto abierto $E \subset R^n$ en R^m se dice que es *continuamente diferenciable* en E si **f** es un mapeo continuo de E en $L(R^n, R^m)$.

Para ser más explícitos, se requiere que a cada $\mathbf{x} \in E$ y a cada $\varepsilon > 0$ corresponda un $\delta > 0$ tal que

$$\|\mathbf{f}'(\mathbf{y}) - \mathbf{f}'(\mathbf{x})\| < \varepsilon$$

si $\mathbf{y} \in E$ y $|\mathbf{x} - \mathbf{y}| < \delta$.

Si esto es así, se dice también que $\mathbf{f}$ es un mapeo- $\mathscr{C}'$, o que $\mathbf{f} \in \mathscr{C}'(E)$.

9.21 Teorema *Supóngase que* $\mathbf{f}$ *mapea un conjunto abierto* $E \subset R^n$ *en* R^m. *Entonces* $\mathbf{f} \in \mathscr{C}'(E)$ *si, y solo si las derivadas parciales* $D_j f_i$ *existan y sean continuas sobre* E *para* $1 \leq i \leq m$, $1 \leq j \leq n$.

Demostración Admítase primero que $\mathbf{f} \in \mathscr{C}'(E)$. De (27) se tiene,

$$(D_j f_i)(\mathbf{x}) = (\mathbf{f}'(\mathbf{x})\mathbf{e}_j) \cdot \mathbf{u}_i$$

para todo i,j, y todo $\mathbf{x} \in E$. Por consiguiente,

$$(D_j f_i)(\mathbf{y}) - (D_j f_i)(\mathbf{x}) = \{[\mathbf{f}'(\mathbf{y}) - \mathbf{f}'(\mathbf{x})]\mathbf{e}_j\} \cdot \mathbf{u}_i$$

y debido a que $|\mathbf{u}_i| = |\mathbf{e}_j| = 1$, se deduce que

$$\begin{aligned}|(D_j f_i)(\mathbf{y}) - (D_j f_i)(\mathbf{x})| &\leq |[\mathbf{f}'(\mathbf{y}) - \mathbf{f}'(\mathbf{x})]\mathbf{e}_j| \\ &\leq \|\mathbf{f}'(\mathbf{y}) - \mathbf{f}'(\mathbf{x})\|.\end{aligned}$$

Por consiguiente $D_j f_i$ es continua.

Para demostrar el inverso, basta considerar el caso $m = 1$. (¿Por qué?) Fijando $\mathbf{x} \in E$ y $\varepsilon > 0$. Ya que E es abierto, hay una bola abierta $S \subset E$, con centro en $\mathbf{x}$ y radio r, y la continuidad de las funciones $D_j f$ muestra que r puede elegirse tal que

(41) $$|(D_j f)(\mathbf{y}) - (D_j f)(\mathbf{x})| < \frac{\varepsilon}{n} \qquad (\mathbf{y} \in S,\ 1 \leq j \leq n).$$

Supóngase que $\mathbf{h} = \Sigma h_j \mathbf{e}_j$, $|\mathbf{h}| < r$, poniendo $\mathbf{v}_0 = \mathbf{0}$, y $\mathbf{v}_k = h_1\mathbf{e}_1 + \cdots + h_k\mathbf{e}_k$, para $1 \leq k \leq n$. Entonces

(42) $$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = \sum_{j=1}^{n} [f(\mathbf{x} + \mathbf{v}_j) - f(\mathbf{x} + \mathbf{v}_{j-1})].$$

Como $|\mathbf{v}_k| < r$ para $1 \leq k \leq n$ y debido a que S es convexo, los segmentos con puntos extremos $\mathbf{x} + \mathbf{v}_{j-1}$ y $\mathbf{x} + \mathbf{v}_j$ están en S. Debido a

que $\mathbf{v}_j = \mathbf{v}_{j-1} + h_j\mathbf{e}_j$, el teorema del valor medio (5.10) muestra que el j-ésimo sumando de (42) es igual a

$$h_j(D_jf)(\mathbf{x} + \mathbf{v}_{j-1} + \theta_j h_j \mathbf{e}_j)$$

para algún $\theta_j \in (0, 1)$, y éste difiere de $h_j(D_jf)(\mathbf{x})$ en menos que $|h_j|\varepsilon/n$, usando (41). De (42) se deduce que

$$\left|f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - \sum_{j=1}^{n} h_j(D_jf)(\mathbf{x})\right| \leq \frac{1}{n}\sum_{j=1}^{n}|h_j|\varepsilon \leq |\mathbf{h}|\varepsilon$$

para todo $\mathbf{h}$ tal que $|\mathbf{h}| < r$.

Esto quiere decir que f es diferenciable en $\mathbf{x}$ y que $f'(\mathbf{x})$ es la función lineal que asigna el número $\Sigma h_j(D_jf)(\mathbf{x})$ al vector $\mathbf{h} = \Sigma h_j\mathbf{e}_j$. La matriz $[f'(\mathbf{x})]$ consta del renglón $(D_1f)(\mathbf{x}), \ldots, (D_nf)(\mathbf{x})$; y ya que $D_1f, \ldots, D_nf$ son funciones continuas sobre E, las observaciones concluyentes de la sección 9.9 muestran que $f \in \mathscr{C}'(E)$.

EL PRINCIPIO DE LA CONTRACCIÓN

Se interrumpirá por ahora el estudio de la diferenciación, para introducir un teorema de punto fijo que es válido en espacios métricos completos arbitrarios. Se usará en la demostración del teorema de la función inversa.

9.22 Definición Sea X un espacio métrico, con métrica d. Si φ mapea X en X y hay un número $c < 1$ tal que

(43) $$d(\varphi(x), \varphi(y)) \leq c\, d(x, y)$$

para todo $x, y \in X$, entonces se dice que φ es una *contracción* de X en X.

9.23 Teorema *Si X es un espacio métrico completo, y si φ es una contracción de X en X, existe entonces un y solo un $x \in X$ tal que $\varphi(x) = x$.*

En otras palabras, φ tiene un único punto fijo. La unicidad es evidente, porque si $\varphi(x) = x$ y $\varphi(y) = y$, entonces (43) nos da $d(x, y) \leq c\, d(x, y)$, lo cual puede suceder sólo cuando $d(x, y) = 0$.

La *existencia* de un punto fijo de φ es la parte esencial del teorema. La demostración proporciona en realidad, un método constructivo para localizar el punto fijo.

Demostración Tómese arbitrariamente un $x_0 \in X$, y defínase $\{x_n\}$ de manera recurrente, haciendo

(44) $$x_{n+1} = \varphi(x_n) \qquad (n = 0, 1, 2, \ldots).$$

Elíjase ahora $c < 1$ de tal suerte que se cumpla (43). Entonces se tiene para $n \geq 1$ que

$$d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq c\, d(x_n, x_{n-1}).$$

Esto da por inducción

(45) $$d(x_{n+1}, x_n) \leq c^n\, d(x_1, x_0) \qquad (n = 0, 1, 2, \ldots).$$

Si $n < m$, se deduce que

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{i=n+1}^{m} d(x_i, x_{i-1}) \\ &\leq (c^n + c^{n+1} + \cdots + c^{m-1})\, d(x_1, x_0) \\ &\leq [(1-c)^{-1}\, d(x_1, x_0)]c^n. \end{aligned}$$

Entonces $\{x_n\}$ es una sucesión de Cauchy. Ya que X es completo, lím $x_n = x$ para algún $x \in X$.

Debido a que φ es una contracción, φ es continuo (de hecho, es uniformemente continuo) sobre X. Por consiguiente

$$\varphi(x) = \lim_{n\to\infty} \varphi(x_n) = \lim_{n\to\infty} x_{n+1} = x.$$

TEOREMA DE LA FUNCIÓN INVERSA

Este teorema establece, vulgarmente hablando, que una aplicación continua diferenciable $\mathbf{f}$ es invertible en una vecindad de cualquier punto $\mathbf{x}$ en el cual es invertible la transformación lineal $\mathbf{f}'(\mathbf{x})$:

9.24 Teorema *Supongamos que* $\mathbf{f}$ *es un mapeo-* $\mathscr{C}'$ *de un conjunto abierto* $E \subset R^n$ *en* R^n, *que* $\mathbf{f}'(\mathbf{a})$ *es invertible para algún* $\mathbf{a} \in E$, *y que* $\mathbf{b} = \mathbf{f}(\mathbf{a})$. *Entonces,*

(*a*) *existen conjuntos abiertos* U *y* V *en* R^n *tales que* $\mathbf{a} \in U$, $\mathbf{b} \in V$ $\mathbf{f}$ *es uno-a-uno en* U *y* $\mathbf{f}(U) = V$;
(*b*) *si* $\mathbf{g}$ *es la inversa de* $\mathbf{f}$ [*que existe, por* (*a*)], *definida en* V *por*

$$\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x} \qquad (\mathbf{x} \in U),$$

entonces, $\mathbf{g} \in \mathscr{C}'(V)$.

Escribiendo la ecuación $\mathbf{y} = \mathbf{f}(\mathbf{x})$ en función de las componentes, llegamos a la siguiente interpretación de la conclusión del teorema: el sistema de n ecuaciones

$$y_i = f_i(x_1, \ldots, x_n) \qquad (1 \leq i \leq n)$$

puede resolverse para $x_1, \ldots, x_n$ en función de $y_1, \ldots, y_n$ si limitamos $\mathbf{x}$ y $\mathbf{y}$ a vecindades suficientemente pequeñas de $\mathbf{a}$ y $\mathbf{b}$; las soluciones son únicas y continuamente diferenciables.

Demostración

(a) Haciendo $\mathbf{f}'(\mathbf{a}) = A$, y eligiendo λ de manera que

(46) $$2\lambda\|A^{-1}\| = 1.$$

Como $\mathbf{f}'$ es continua en $\mathbf{a}$, hay una bola abierta $U \subset E$, con centro en $\mathbf{a}$, tal que

(47) $$\|\mathbf{f}'(\mathbf{x}) - A\| < \lambda \qquad (\mathbf{x} \in U).$$

Asóciese a cada $\mathbf{y} \in R^n$ una función φ, definida por

(48) $$\varphi(\mathbf{x}) = \mathbf{x} + A^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x})) \qquad (\mathbf{x} \in E).$$

Nótese que $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ *si, y solo si* $\mathbf{x}$ *es un punto fijo de* φ.

Como $\varphi'(\mathbf{x}) = I - A^{-1}\mathbf{f}'(\mathbf{x}) = A^{-1}(A - \mathbf{f}'(\mathbf{x}))$, (46) y (47) implican que

(49) $$\|\varphi'(\mathbf{x})\| < \tfrac{1}{2} \qquad (\mathbf{x} \in U).$$

De aquí que, por el Teorema 9.19

(50) $$|\varphi(\mathbf{x}_1) - \varphi(\mathbf{x}_2)| \leq \tfrac{1}{2}|\mathbf{x}_1 - \mathbf{x}_2| \qquad (\mathbf{x}_1, \mathbf{x}_2 \in U),$$

Se deduce entonces que φ tiene a lo más un punto fijo en U, así que $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ a lo más para un $\mathbf{x} \in U$.

Por esto $\mathbf{f}$ *es* 1-1 *en* U.

En seguida se pone $V = \mathbf{f}(U)$, y se elige $\mathbf{y}_0 \in V$. Entonces para algún $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ para algún $\mathbf{x}_0 \in U$. Sea B una bola abierta con centro en $\mathbf{x}_0$ y radio $r > 0$, tan pequeño que su cerradura $\overline{B}$ está en U.

Se mostrará que $\mathbf{y} \in V$ siempre y cuando $|\mathbf{y} - \mathbf{y}| < \lambda r$. Esto demuestra, por supuesto, que V es abierto.

Si se fija $\mathbf{y}$, $|\mathbf{y} - \mathbf{y}_0| < \lambda r$, y con φ como en (48)

$$|\varphi(\mathbf{x}_0) - \mathbf{x}_0| = |A^{-1}(\mathbf{y} - \mathbf{y}_0)| < \|A^{-1}\|\lambda r = \frac{r}{2}.$$

Si $\mathbf{x} \in B$, de (50) se deduce por tanto, que

$$\begin{aligned}|\varphi(\mathbf{x}) - \mathbf{x}_0| &\le |\varphi(\mathbf{x}) - \varphi(\mathbf{x}_0)| + |\varphi(\mathbf{x}_0) - \mathbf{x}_0| \\ &< \frac{1}{2}|\mathbf{x} - \mathbf{x}_0| + \frac{r}{2} \le r;\end{aligned}$$

en consecuencia $\varphi(\mathbf{x}) \in B$. Nótese que (50) se cumple si $\mathbf{x}_1 \in \overline{B}$, $\mathbf{x}_2 \in \overline{B}$.

Entonces φ es una contracción de $\overline{B}$ en $\overline{B}$. Siendo un subconjunto cerrado de R^n es completo. Por tanto, el Teorema 9.23 implica que φ tiene un punto fijo $\mathbf{x} \in \overline{B}$. Para este $\mathbf{x}$, $f(\mathbf{x}) = \mathbf{y}$. De aquí que $\mathbf{y} \in (\overline{B}) \subset \mathbf{f}(U) = V$.

Esto demuestra la parte (a) del teorema.

(b) Elíjanse $\mathbf{y} \in V$, $\mathbf{y} + \mathbf{k} \in V$. Entonces existe $\mathbf{x} \in U$, $\mathbf{x} + \mathbf{h} \in U$, de tal manera que $\mathbf{y} = \mathbf{f}(\mathbf{x})$, $\mathbf{y} + \mathbf{k} = \mathbf{f}(\mathbf{x} + \mathbf{h})$. Con φ como en (48) se tiene

$$\varphi(\mathbf{x} + \mathbf{h}) - \varphi(\mathbf{x}) = \mathbf{h} + A^{-1}[\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x} + \mathbf{h})] = \mathbf{h} - A^{-1}\mathbf{k}.$$

Por (50), $|\mathbf{h} - A^{-1}\mathbf{k}| \le \frac{1}{2}|\mathbf{h}|$. Por esto $|A^{-1}\mathbf{k}| \ge \frac{1}{2}|\mathbf{h}|$, y

(51) $$|\mathbf{h}| \le 2\|A^{-1}\|\,|\mathbf{k}| = \lambda^{-1}|\mathbf{k}|.$$

Por (46), (47) y el Teorema 9.8, $\mathbf{f}'(\mathbf{x})$ tiene una inversa, digamos T. Como

$$\mathbf{g}(\mathbf{y} + \mathbf{k}) - \mathbf{g}(\mathbf{y}) - T\mathbf{k} = \mathbf{h} - T\mathbf{k} = -T[\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}],$$

(51) implica que

$$\frac{|\mathbf{g}(\mathbf{y} + \mathbf{k}) - \mathbf{g}(\mathbf{y}) - T\mathbf{k}|}{|\mathbf{k}|} \le \frac{\|T\|}{\lambda} \cdot \frac{|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}|}{|\mathbf{h}|}.$$

Cuando $\mathbf{k} \to \mathbf{0}$, de (51) se ve que $\mathbf{h} \to \mathbf{0}$. Entonces, el miembro derecho de la desigualdad anterior tiende hacia 0. Por consiguier e le

pasa lo mismo al izquierdo. Se ha demostrado entonces que $\mathbf{g}'(\mathbf{y}) = T$. Pero como se eligió T como el inverso de $\mathbf{f}'(\mathbf{x}) = \mathbf{f}'(\mathbf{g}(\mathbf{y}))$. Entonces

(52) $$\mathbf{g}'(\mathbf{y}) = \{\mathbf{f}'(\mathbf{g}(\mathbf{y}))\}^{-1} \qquad (\mathbf{y} \in V).$$

Nótese finalmente, que $\mathbf{g}$ es un mapeo continuo de V sobre U (porque $\mathbf{g}$ es diferenciable), que $\mathbf{f}$ es un mapeo continuo de U en el conjunto Ω de todos los elementos invertibles de $L(R^n)$, y que la inversión es un mapeo continuo de Ω sobre Ω, debido al Teorema 9.8. Si se combinan estos efectos con (52), se ve que $\mathbf{g} \in \mathscr{C}'(V)$.

Esto completa la demostración.

Observación La fuerza total de la suposición de que $\mathbf{f} \in \mathscr{C}'(E)$ sólo se usó en el último párrafo de la demostración anterior. Todo lo posterior a la ecuación (52) se dedujo a partir de la existencia de $\mathbf{f}'(\mathbf{x})$ para $\mathbf{x} \in E$, la invertibilidad de $\mathbf{f}'(\mathbf{a})$, y la continuidad de $\mathbf{f}'$ en el punto $\mathbf{a}$. El artículo de A. Nijenhuis publicado en la *Amer. Math. Monthly*, tomo 81, 1974, páginas 969 a 980, se refiere a esto.

Como consecuencia inmediata del apartado (*a*) del teorema de la función inversa, podemos enunciar:

9.25 Teorema *Si* $\mathbf{f}$ *es un mapeo-* $\mathscr{C}'$ *de un conjunto abierto* $E \subset R^n$ *en* R^n *y si* $\mathbf{f}'(\mathbf{x})$ *es invertible para todo* $\mathbf{x} \in E$, *entonces* $\mathbf{f}(W)$ *es un subconjunto abierto de* R^n *para todo conjunto abierto* $W \subset E$.

En otras palabras, $\mathbf{f}$ es un *mapeo abierto* de E en R^n.

Las hipótesis hechas en este teorema aseguran que cada punto $\mathbf{x} \in E$ tiene una vecindad en la que $\mathbf{f}$ es uno-a-uno, lo que puede expresarse diciendo que $\mathbf{f}$ es uno-a-uno *localmente* en E, pero $\mathbf{f}$ no es necesariamente 1-1 en E en tales circunstancias. Como ejemplo, véase el Ejercicio 17.

EL TEOREMA DE LA FUNCIÓN IMPLÍCITA

Si f es una función real continuamente diferenciable en el plano, entonces $f(x, y) = 0$ puede resolverse para y en términos de x en una vecindad de cualquier punto (a,b) en el cual $f(a,b) = 0$ y $\partial f/\partial y \neq 0$. De la misma manera, se puede resolver para x en términos de y cerca de (a,b) si $\partial f/\partial x \neq 0$ en (a,b). Como un ejemplo sencillo que ilustra la necesidad de suponer que $\partial f/\partial y \neq 0$ se tiene $f(x, y) = x^2 + y^2 - 1$.

La proposición anterior, por cierto bastante informal, es el caso más simple (el caso $m = n = 1$ del Teorema 9.28) del llamado "teorema de la función implícita". Su demostración se basa fuertemente en el hecho de que las transformaciones continuamente diferenciables en la mayoría de los casos se

comportan localmente como sus derivadas. De acuerdo a esto, primero se demostrará el Teorema 9.27, que es la versión lineal del Teorema 9.28.

9.26 Notación Si $\mathbf{x} = (x_1, \ldots, x_n) \in R^n$ y $\mathbf{y} = (y_1, \ldots, y_m) \in R^m$, se escribirá $(\mathbf{x}, \mathbf{y})$ en lugar del punto (o vector)

$$(x_1, \ldots, x_n, y_1, \ldots, y_m) \in R^{n+m}.$$

De aquí en adelante, el primer elemento en $(\mathbf{x}, \mathbf{y})$ o en un símbolo similar, será siempre un vector en R^n, y el segundo será un vector en R^m.

Cada $A \in L(R^{n+m}, R^n)$ puede separarse en dos transformaciones lineales A_x y A_y, definidas por

$$A_x\mathbf{h} = A(\mathbf{h}, \mathbf{0}), \qquad A_y\mathbf{k} = A(\mathbf{0}, \mathbf{k}) \tag{53}$$

para cualquier $\mathbf{h} \in R^n$, $\mathbf{k} \in R^m$. Entonces $A_x \in L(R^n)$, $A_y \in L(R^m, R^n)$, y

$$A(\mathbf{h}, \mathbf{k}) = A_x\mathbf{h} + A_y\mathbf{k}. \tag{54}$$

La versión lineal del teorema de la función implícita es ahora casi obvia.

9.27 Teorema *Si $A \in L(R^{n+m}, R^m)$ y si A_x es invertible, entonces a cada* $\mathbf{k} \in R^m$ *le corresponde un* $\mathbf{h} \in R^n$ *único, y tal que* $A(\mathbf{h}, \mathbf{k}) = 0$.

Este $\mathbf{h}$ *puede calcularse a partir de* $\mathbf{k}$ *por medio de la fórmula*

$$\mathbf{h} = -(A_x)^{-1}A_y\mathbf{k}. \tag{55}$$

Demostración De (54), $A(\mathbf{h},\mathbf{k}) = \mathbf{0}$ si, y solo si

$$A_x\mathbf{h} + A_y\mathbf{k} = \mathbf{0},$$

que es la misma que (55) cuando A_x es invertible.

En otras palabras, la conclusión del Teorema 9.27 es que la ecuación $A = (\mathbf{h},\mathbf{k}) = \mathbf{0}$ puede resolverse (de manera única) para $\mathbf{h}$ si $\mathbf{k}$ se conoce, y que la solución $\mathbf{h}$ es una función lineal de $\mathbf{k}$. Aquellos lectores que estén familiarizados con el álgebra lineal, podrán darse cuenta que esta es una proposición sobre sistemas de ecuaciones lineales muy conocida.

9.28 Teorema *Sea* $\mathbf{f}$ *un mapeo-* $\mathscr{C}'$ *de un conjunto abierto* $E \subset R^{n+m}$ *en* R^m, *tal que* $\mathbf{f}(\mathbf{a},\mathbf{b}) = \mathbf{0}$ *para algún punto* $(\mathbf{a},\mathbf{b}) \in E$.

Se pone $A = \mathbf{f}'(\mathbf{a},\mathbf{b})$ *y se supone que* A_x *es invertible.*

Entonces existen conjuntos abiertos $U \subset R^{n+m}$ *y* $W \subset R^m$, *con* $(\mathbf{a},\mathbf{b}) \in U$ *y* $\mathbf{b} \in W$, *que tienen la siguiente propiedad:*

A cada **y** $\in$ *W le corresponde un* **x** *único, tal que*

(56) $$(\mathbf{x}, \mathbf{y}) \in U \qquad \text{y} \qquad \mathbf{f}(\mathbf{x}, \mathbf{y}) = 0.$$

Si esta **x** *se define como* **g(y)**, *entonces* **g** *es un mapeo* $\mathscr{C}'$ *de W en* R^n, **g(b)** = **a**,

(57) $$\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \mathbf{0} \qquad (\mathbf{y} \in W),$$

y

(58) $$\mathbf{g}'(\mathbf{b}) = -(A_x)^{-1}A_y.$$

La función *g* está definida "implícitamente" por (57). De aquí el nombre del teorema.

La ecuación **f(x,y)** = **0** puede escribirse como un sistema de *n* ecuaciones con $n + m$ variables:

(59) $$\begin{array}{l} f_1(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_n(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0. \end{array}$$

La suposición de que A_x es invertible significa que la matriz *n* por *n*

$$\begin{bmatrix} D_1f_1 & \cdots & D_nf_1 \\ \cdots & \cdots & \cdots \\ D_1f_n & \cdots & D_nf_n \end{bmatrix}$$

evaluada en **(a,b)** define un operador lineal invertible en R^n; dicho de otra manera, sus vectores columna deben ser independientes, o de manera equivalente, su determinante debe ser $\neq 0$. (Véase el Teorema 9.36.) Si además, (59) se cumple cuando **x** = **a** y **y** = **b**, entonces la conclusión del teorema es que (59) puede ser resuelta para $x_1, \ldots, x_n$ en términos de $y_1, \ldots, y_m$, para cada **y** cercana a **b**, y que estas soluciones son funciones continuamente diferenciables de **y**.

Demostración Se define **F** por medio de

(60) $$\mathbf{F}(\mathbf{x}, \mathbf{y}) = (\mathbf{f}(\mathbf{x}, \mathbf{y}), \mathbf{y}) \qquad ((\mathbf{x}, \mathbf{y}) \in E).$$

Entonces **F** es un mapeo-$\mathscr{C}'$ de *E* en R^{n+m}. Se pide que **f**′**(a,b)** sea un elemento invertible de $L(R^{n+m})$:

Como **f(a,b)** = **0**, se tiene que

$$\mathbf{f}(\mathbf{a} + \mathbf{h}, \mathbf{b} + \mathbf{k}) = A(\mathbf{h}, \mathbf{k}) + \mathbf{r}(\mathbf{h}, \mathbf{k}),$$

en donde **r** es el residuo que aparece en la definición de $\mathbf{f}'(\mathbf{a},\mathbf{b})$. Debido a que

$$\begin{aligned}\mathbf{F}(\mathbf{a}+\mathbf{h}, \mathbf{b}+\mathbf{k}) - \mathbf{F}(\mathbf{a}, \mathbf{b}) &= (\mathbf{f}(\mathbf{a}+\mathbf{h}, \mathbf{b}+\mathbf{k}), \mathbf{k}) \\ &= (\mathbf{A}(\mathbf{h}, \mathbf{k}), \mathbf{k}) + (\mathbf{r}(\mathbf{h}, \mathbf{k}), \mathbf{0})\end{aligned}$$

se deduce que $\mathbf{F}'(\mathbf{a},\mathbf{b})$ es el operador lineal sobre R^{n+m} que mapea $(\mathbf{h},\mathbf{k})$ en $(A(\mathbf{h},\mathbf{k}), \mathbf{k})$. Si este vector imagen es $\mathbf{0}$, entonces $A(\mathbf{h},\mathbf{k}) = \mathbf{0}$ y $\mathbf{k} = \mathbf{0}$, de aquí que $A(\mathbf{h},\mathbf{0}) = \mathbf{0}$, y el Teorema 9.27 implica que $\mathbf{h} = \mathbf{0}$. Y se deduce que $\mathbf{F}'(\mathbf{a},\mathbf{b})$ es 1-1; en consecuencia es invertible (Teorema 9.5).

El teorema de la función inversa puede aplicarse por tanto a **F**. Muestra que existen conjuntos abiertos U y V en R^{n+m}, con $(\mathbf{a},\mathbf{b}) \in U$, $(\mathbf{0},\mathbf{b}) \in V$, tales que **F** es un mapeo 1-1 de U sobre V.

Sea ahora W el conjunto de todos los $\mathbf{y} \in R^m$ tales que $(\mathbf{0},\mathbf{y}) \in V$. Nótese que $\mathbf{b} \in W$.

Es claro que W es abierto porque V es abierto.

Si $\mathbf{y} \in W$, entonces $(\mathbf{0},\mathbf{y}) = \mathbf{F}(\mathbf{x},\mathbf{y})$ para algún $(\mathbf{x},\mathbf{y}) \in U$. Por (60), $\mathbf{f}(\mathbf{x},\mathbf{y}) = \mathbf{0}$ para este **x**.

Con el mismo **y**, supóngase que $(\mathbf{x}',\mathbf{y}) \in U$ y $\mathbf{f}(\mathbf{X},\mathbf{y}) = \mathbf{0}$. Entonces

$$\mathbf{F}(\mathbf{x}', \mathbf{y}) = (\mathbf{f}(\mathbf{x}', \mathbf{y}), \mathbf{y}) = (\mathbf{f}(\mathbf{x}, \mathbf{y}), \mathbf{y}) = \mathbf{F}(\mathbf{x}, \mathbf{y}).$$

Como **F** es 1-1 en U, se deduce que $\mathbf{x}' = \mathbf{x}$.

Esto demuestra la primera parte del teorema.

Para demostrar la segunda parte, defínase $\mathbf{g}(\mathbf{y})$, para $\mathbf{y} \in W$, de manera que $(\mathbf{g}(\mathbf{y}), \mathbf{y}) \in U$ y (57) se cumpla. Entonces

(61) $$\mathbf{F}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = (\mathbf{0}, \mathbf{y}) \qquad (\mathbf{y} \in W).$$

Si **G** es el mapeo de V sobre U que invierte **F**, entonces por el teorema de la función inversa $\mathbf{G} \in \mathscr{C}'$, y (61) nos da

(62) $$(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \mathbf{G}(\mathbf{0}, \mathbf{y}) \qquad (\mathbf{y} \in W).$$

Como $\mathbf{G} \in \mathscr{C}'$, de (62) se ve que $\mathbf{g} \in \mathscr{C}'$.

Finalmente, para calcular $\mathbf{g}'(\mathbf{b})$, se hace $\mathbf{g}(\mathbf{y}), \mathbf{y}) = \Phi(\mathbf{y})$. Entonces

(63) $$\Phi'(\mathbf{y})\mathbf{k} = (\mathbf{g}'(\mathbf{y})\mathbf{k}, \mathbf{k}) \qquad (\mathbf{y} \in W, \mathbf{k} \in R^m).$$

Por (57), $\mathbf{f}(\Phi(\mathbf{y})) = \mathbf{0}$ en W. Por lo tanto, la regla de la cadena muestra que

$$\mathbf{f}'(\Phi(\mathbf{y}))\Phi'(\mathbf{y}) = 0.$$

Cuando $\mathbf{y} = \mathbf{b}$, entonces $\Phi(\mathbf{y}) = (\mathbf{a},\mathbf{b})$, y $\mathbf{f}'(\Phi(\mathbf{y})) = A$. Por lo anterior,

(64) $$A\Phi'(\mathbf{b}) = 0.$$

Ahora se deduce de (64), (63) y (54) que

$$A_x\mathbf{g}'(\mathbf{b})\mathbf{k} + A_y\mathbf{k} = A(\mathbf{g}'(\mathbf{b})\mathbf{k}, \mathbf{k}) = A\Phi'(\mathbf{b})\mathbf{k} = \mathbf{0}$$

para cada $\mathbf{k} \in R^m$. Entonces

(65) $$A_x\mathbf{g}'(\mathbf{b}) + A_y = 0.$$

Esto equivale a (58), y la demostración queda completa.

Nota. En términos de las componentes de $\mathbf{f}$ y $\mathbf{g}$, (65) se convierte en

$$\sum_{j=1}^{n} (D_j f_i)(\mathbf{a}, \mathbf{b})(D_k g_j)(\mathbf{b}) = -(D_{n+k} f_i)(\mathbf{a}, \mathbf{b})$$

o

$$\sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j}\right)\left(\frac{\partial g_j}{\partial y_k}\right) = -\left(\frac{\partial f_i}{\partial y_k}\right)$$

en donde $1 \le i \le n$, $1 \le k \le m$.

Para cada k, este es un sistema de n ecuaciones lineales en las cuales las derivadas $\partial g_j/\partial y_k$ $(1 \le j \le n)$ son las incógnitas.

9.29 Ejemplo Tomando $n = 2$, $m = 3$, y considerando el mapeo $\mathbf{f} = (f_1, f_2)$ de R^5 en R^2 dado por

$$f_1(x_1, x_2, y_1, y_2, y_3) = 2e^{x_1} + x_2 y_1 - 4y_2 + 3$$
$$f_2(x_1, x_2, y_1, y_2, y_3) = x_2 \cos x_1 - 6x_1 + 2y_1 - y_3 .$$

Si $\mathbf{a} = (0, 1)$ y $\mathbf{b} = (3, 2, 7)$, entonces $\mathbf{f}(\mathbf{a},\mathbf{b}) = 0$.

Con respecto a la base estándar, la matriz de la transformación $A = \mathbf{f}'(\mathbf{a},\mathbf{b})$ es

$$[A] = \begin{bmatrix} 2 & 3 & 1 & -4 & 0 \\ -6 & 1 & 2 & 0 & -1 \end{bmatrix}.$$

En consecuencia

$$[A_x] = \begin{bmatrix} 2 & 3 \\ -6 & 1 \end{bmatrix}, \qquad [A_y] = \begin{bmatrix} 1 & -4 & 0 \\ 2 & 0 & -1 \end{bmatrix}.$$

Se ve que los vectores columna de $[A_x]$ son independientes. De aquí que A_x es invertible y el teorema de la función implícita asegura la existencia de un mapeo- $\mathscr{C}'$**g**, definido en una vecindad de (3, 2, 7), tal que **g**(3, 2, 7) = (0, 1) y **f**(**g**(**y**), **y**) = **0**.

Puede usarse (58) para calcular **g**'(3, 2, 7): Ya que

$$[(A_x)^{-1}] = [A_x]^{-1} = \frac{1}{20}\begin{bmatrix} 1 & -3 \\ 6 & 2 \end{bmatrix}$$

(58) nos da

$$[\mathbf{g}'(3,2,7)] = -\frac{1}{20}\begin{bmatrix} 1 & -3 \\ 6 & 2 \end{bmatrix}\begin{bmatrix} 1 & -4 & 0 \\ 2 & 0 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{5} & -\frac{3}{20} \\ -\frac{1}{2} & \frac{6}{5} & \frac{1}{10} \end{bmatrix}.$$

En términos de derivadas parciales, la conclusión es que

$$\begin{array}{lll} D_1 g_1 = \frac{1}{4} & D_2 g_1 = \frac{1}{5} & D_3 g_1 = -\frac{3}{20} \\ D_1 g_2 = -\frac{1}{2} & D_2 g_2 = \frac{6}{5} & D_3 g_2 = \frac{1}{10} \end{array}$$

en el punto (3, 2, 7).

EL TEOREMA DEL RANGO

A pesar de que este teorema no es tan importante como los teoremas de la función inversa y de la función implícita, se incluye como otra ilustración interesante del principio general que expresa que el comportamiento local de un mapeo **F** continuamente diferenciable cerca de un punto **x** es similar al de la transformación lineal **F**'(**x**).

Antes de establecerlo se necesita más información acerca de las transformaciones lineales.

9.30 Definiciones Supóngase que X y Y son espacios vectoriales, y que $A \in L(X, Y)$, de la misma forma que en la Definición 9.6. El *espacio nulo* de A, $\mathscr{N}(A)$, es el conjunto de todos los $\mathbf{x} \in X$ en los cuales $A\mathbf{x} = \mathbf{0}$. Es claro que $\mathscr{N}(A)$ es un espacio vectorial en X.

De la misma forma, el *rango* de A, $\mathscr{R}(A)$, es un espacio vectorial en Y.

El *rango* de A se define como la dimensión de $\mathscr{R}(A)$.

Por ejemplo, los elementos invertibles de $L(R^n)$ son precisamente aquellos cuyo rango es n. Esto se deduce a partir del Teorema 9.5.

Si $A \in L(X, Y)$ y A tiene rango 0, entonces $A\mathbf{x} = \mathbf{0}$ para todo $x \in A$, y de aquí $\mathscr{N}(A) = X$. Véase el Ejercicio 25, relacionado con esto.

9.31 Proyecciones Sea X un espacio vectorial. Se dice que un operador $P \in L(X)$ es una *proyección en X* si $P^2 = P$.

Para ser más explícitos, el requisito es que $P(P\mathbf{x}) = P\mathbf{x}$ para cada $\mathbf{x} \in X$. Dicho de otro modo, P fija cada vector en su "rango" (P).

A continuación se tienen algunas propiedades elementales de las proyecciones:

(*a*) *Si P es una proyección en X, entonces cada $\mathbf{x} \in X$ tiene una representación única de la forma*

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$$

donde $\mathbf{x}_1 \in \mathscr{R}(P)$, $\mathbf{x}_2 \in \mathscr{N}(P)$.

Para obtener la representación, póngase $\mathbf{x}_1 = P\mathbf{x}$, $\mathbf{x}_2 = \mathbf{x} - \mathbf{x}_1$. Entonces $P\mathbf{x}_2 = P\mathbf{x} - P\mathbf{x}_1 = P\mathbf{x} - P^2\mathbf{x} = \mathbf{0}$. Por lo que se refiere a la unicidad, aplíquese P a la ecuación $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$. Como $\mathbf{x}_1 \in \mathscr{R}(P)$, $P\mathbf{x}_1 = \mathbf{x}_1$; debido a que $P\mathbf{x}_2 = \mathbf{0}$, se deduce que $\mathbf{x}_1 = P\mathbf{x}$.

(*b*) *Si X es un espacio vectorial de dimensión finita y si X es un espacio vectorial en X, entonces hay una proyección P en X con $\mathscr{R}(P) = X_1$.*

Si X_1 contiene solo al $\mathbf{0}$, esto es trivial: se pone $P\mathbf{x} = \mathbf{0}$ para todo $\mathbf{x} \in X$.

Admitiendo que dim $X_1 = k > 0$. Por el Teorema 9.3, X tiene entonces una base $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ tal que $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ es una base de X_1. Si se define

$$P(c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n) = c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k$$

para escalares $c_1, \ldots, c_n$.

Entonces $P\mathbf{x} = \mathbf{x}$ para cada $\mathbf{x} \in X_1$, y $X_1 = \mathscr{R}(P)$.

Nótese que $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ es una base de $\mathscr{N}(P)$. Nótese también que hay una infinidad de proyecciones en X, con "rango" X_1, si $0 < \dim X_1 < \dim X$.

9.32 Teorema *Supóngase que m, n, r son enteros no negativos, $m \geq r$, $n \geq r$, $\mathbf{F}$ en un mapeo-$\mathscr{C}'$ de un conjunto abierto $E \subset R^n$ en R^m, y $\mathbf{F}'(\mathbf{x})$ tiene rango r para cada $\mathbf{x} \in E$.*

Fíjese $\mathbf{a} \in E$, y póngase $A = \mathbf{F}(\mathbf{a})$, ahora sea Y_1 el "rango" de A, y P una proyección en R^m cuyo "rango" es Y_1. Sea Y_2 el espacio nulo de P.

Entonces hay conjuntos abiertos U y V, con $\mathbf{a} \in U$, $U \subset E$, y un mapeo-$\mathscr{C}'$ $\mathbf{H}$, 1-1 de V sobre U (cuyo inverso es también de clase $\mathscr{C}'$) tal que

$$\mathbf{F}(\mathbf{H}(\mathbf{x})) = A\mathbf{x} + \varphi(A\mathbf{x}) \qquad (\mathbf{x} \in V) \tag{66}$$

donde φ es un mapeo-$\mathscr{C}'$ de un conjunto abierto $A(V) \subset Y_1$ en Y_2.

Después de la demostración se dará una descripción más geométrica de la información que contiene (66).

Demostración Si $r = 0$, el Teorema 9.19 muestra que $\mathbf{F}(\mathbf{x})$ es constante en una vecindad U de $\mathbf{a}$, y (66) evidentemente se cumple, con $V = U$, $\mathbf{H}(\mathbf{x}) = \mathbf{x}$, $\varphi(\mathbf{0}) = \mathbf{F}(\mathbf{a})$.

De aquí en adelante se supondrá $r > 0$. Como dim $Y_1 = r$, Y_1 tiene una base $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$. Se elige $\mathbf{z}_i \in R^n$ de manera que $A\mathbf{z}_i = \mathbf{y}_i (1 \leq i \leq r)$, y se define un mapeo lineal S de Y_1 en R^n de la siguiente manera

(67) $$S(c_1\mathbf{y}_1 + \cdots + c_r\mathbf{y}_r) = c_1\mathbf{z}_1 + \cdots + c_r\mathbf{z}_r$$

para todos los escalares $c_1, \ldots, c_r$.

Entonces $AS\mathbf{y}_i = A\mathbf{z}_i = \mathbf{y}_i$ para $1 \leq i \leq r$. De aquí se tiene

(68) $$AS\mathbf{y} = \mathbf{y} \qquad (\mathbf{y} \in Y_1).$$

Defínase un mapeo $\mathbf{G}$ de E en R^n haciendo

(69) $$\mathbf{G}(\mathbf{x}) = \mathbf{x} + SP[\mathbf{F}(\mathbf{x}) - A\mathbf{x}] \qquad (\mathbf{x} \in E).$$

Como $\mathbf{F}'(\mathbf{a}) = A$, la diferenciación de (69) muestra que $\mathbf{G}'(\mathbf{a}) = I$, el operador identidad sobre R^n. Por el teorema de la función inversa, hay conjuntos abiertos U y V en R^n, con $\mathbf{a} \in U$, de tal manera que $\mathbf{G}$ es un mapeo 1-1 de U sobre V cuyo inverso $\mathbf{H}$ es también de clase $\mathscr{C}'$. Además, disminuyendo U y V, si fuera necesario, se puede ordenar de tal forma que V sea convexo y $\mathbf{H}'(\mathbf{x})$ invertible para cada $\mathbf{x} \in V$.

Nótese que $ASPA = A$, ya que $PA = A$ y (68) se cumple. Por tanto (69) nos da

(70) $$A\mathbf{G}(\mathbf{x}) = P\mathbf{F}(\mathbf{x}) \qquad (\mathbf{x} \in E).$$

En particular, se cumple (70) para $\mathbf{x} \in U$. Si se reemplaza $\mathbf{x}$ por $\mathbf{H}(\mathbf{x})$, se obtiene

(71) $$P\mathbf{F}(\mathbf{H}(\mathbf{x})) = A\mathbf{x} \qquad (\mathbf{x} \in V).$$

Ahora se define

(72) $$\psi(\mathbf{x}) = \mathbf{F}(\mathbf{H}(\mathbf{x})) - A\mathbf{x} \qquad (\mathbf{x} \in V).$$

Como $PA = A$, (71) implica que $P\psi(\mathbf{x}) = \mathbf{0}$ para todo $\mathbf{x} \in V$. De esto ψ es un mapeo- $\mathscr{C}'$ de V en Y_2.

Como V es abierto, es evidente que $A(V)$ es un subconjunto abierto de su "rango" $\mathscr{R}(A) = Y_1$.

Para completar la demostración, es decir, pasar de (72) a (66), tiene que mostrarse que hay un mapeo-$\mathscr{C}'$ φ de $A(V)$ en Y_2 que satisface

(73) $$\varphi(A\mathbf{x}) = \psi(\mathbf{x}) \qquad (\mathbf{x} \in V).$$

Como un paso hacia (73), primero se probará que

(74) $$\psi(\mathbf{x}_1) = \psi(\mathbf{x}_2)$$

si $\mathbf{x}_1 \in V$, $\mathbf{x}_2 \in V$, $A\mathbf{x}_1 = A\mathbf{x}_2$.

Poniendo $\Phi(\mathbf{x}) = \mathbf{F}(\mathbf{H}(\mathbf{x}))$, para $\mathbf{x} \in V$. Como $\mathbf{H}'(\mathbf{x})$ tiene rango n para cada $\mathbf{x} \in V$, y $\mathbf{F}'(\mathbf{x})$ tiene rango r para cada $\mathbf{x} \in U$, se deduce

(75) $$\text{rango } \Phi'(\mathbf{x}) = \text{rango } \mathbf{F}'(\mathbf{H}(\mathbf{x}))\mathbf{H}'(\mathbf{x}) = r \qquad (\mathbf{x} \in V).$$

Con $\mathbf{x} \in V$ fijo, sea M el "rango" de $\Phi(\mathbf{x})$. Entonces $M \subset R^m$, dim $M = r$. Por (71),

(76) $$P\Phi'(\mathbf{x}) = A.$$

Por lo anterior P mapea M sobre $\mathscr{R}(A) = Y_1$. Ya que M y Y_1 tienen la misma dimensión, se deduce que P (restringido a M) es 1-1.

Supóngase ahora que $A\mathbf{h} = \mathbf{0}$. Entonces de (76) $P\Phi'(\mathbf{x})\mathbf{h} = \mathbf{0}$. Pero $\Phi'(\mathbf{x})\mathbf{h} \in M$, y P es 1-1 sobre M. En consecuencia, $\Phi'(\mathbf{x})\mathbf{h} = \mathbf{0}$. Observando (72) se ve que se ha probado lo siguiente:

Si $\mathbf{x} \in V$ *y* $A\mathbf{h} = \mathbf{0}$, *entonces* $\psi'(\mathbf{x})\mathbf{h} = \mathbf{0}$.

Se puede probar ahora (74). Suponiendo que $\mathbf{x}_1 \in V$, $\mathbf{x}_2 \in V$, $A\mathbf{x}_1 = A\mathbf{x}_2$. Poniendo $\mathbf{h} = \mathbf{x}_2 - \mathbf{x}_1$ y definiendo

(77) $$\mathbf{g}(t) = \psi(\mathbf{x}_1 + t\mathbf{h}) \qquad (0 \le t \le 1).$$

La convexibilidad de V muestra que $\mathbf{x}_1 + t\mathbf{h} \in V$ para estos t. En consecuencia,

(78) $$\mathbf{g}'(t) = \psi'(\mathbf{x}_1 + t\mathbf{h})\mathbf{h} = \mathbf{0} \qquad (0 \le t \le 1),$$

así que $\mathbf{g}(1) = \mathbf{g}(0)$. Pero $\mathbf{g}(1) = \psi(\mathbf{x}_2)$ y $\mathbf{g}(0) = \psi(\mathbf{x}_1)$. Esto demuestra (74).

De (74), se ve que $\psi(\mathbf{x})$ depende sólo de $A\mathbf{x}$, para $\mathbf{x} \in V$. De lo anterior, (73) define φ sin ambigüedad en $A(V)$. Sólo resta probar que $\varphi \in \mathscr{C}'$.

Con $\mathbf{y}_0 \in A(V)$ fijo y $\mathbf{x}_0 \in V$ *fijo de manera que* $A\,\mathbf{x}_0 = \mathbf{y}_0$. Como V es abierto, $\mathbf{y}_0$ tiene una vecindad W en Y_1 tal que el vector

(79) $$\mathbf{x} = \mathbf{x}_0 + S(\mathbf{y} - \mathbf{y}_0)$$

está en V para todo $\mathbf{y} \in W$. Por (68) se tiene,

$$A\mathbf{x} = A\mathbf{x}_0 + \mathbf{y} - \mathbf{y}_0 = \mathbf{y}.$$

Entonces (73) y (79) nos dan

(80) $$\varphi(\mathbf{y}) = \psi(\mathbf{x}_0 - S\mathbf{y}_0 + S\mathbf{y}) \qquad (\mathbf{y} \in W).$$

Esta fórmula muestra que $\varphi \in \mathscr{C}'$ en W, en consecuencia en $A(V)$, debido a que $\mathbf{y}_0$ se escogió arbitrariamente en $A(V)$.

La demostración está completa ahora.

Esto es lo que el teorema dice acerca de la geometría del mapeo $\mathbf{F}$.

Si $\mathbf{y} \in \mathbf{F}(U)$, entonces $\mathbf{y} = \mathbf{F}(\mathbf{H}(\mathbf{x}))$ para algún $\mathbf{x} \in V$, y (66) muestra que $P\mathbf{y} = A\mathbf{x}$. Por tanto,

(81) $$\mathbf{y} = P\mathbf{y} + \varphi(P\mathbf{y}) \qquad (\mathbf{y} \in \mathbf{F}(U)).$$

Esto muestra que $\mathbf{y}$ está determinado por su proyección $P\mathbf{y}$, y que P, restringido a $\mathbf{F}(U)$, es un mapeo 1-1 de $\mathbf{F}(U)$ sobre $A(V)$. Entonces $\mathbf{F}(U)$ es una "superficie de dimensión-r" con un punto "sobre" cada punto de $A(V)$. Puede también considerarse a $\mathbf{F}(U)$ como la gráfica de φ.

Si como en la demostración $\Phi(\mathbf{x}) = \mathbf{F}(\mathbf{H}(\mathbf{x}))$, entonces (66) muestra que los conjuntos de nivel de Φ (estos son los conjuntos sobre los cuales Φ alcanza un valor dado) son precisamente los conjuntos de nivel de A en V. Estos son "planos" porque son intersecciones con V de translaciones del espacio vectorial $\mathscr{N}(A)$. Nótese que $\dim \mathscr{N}(A) = n - r$ (Ejercicio 25).

Los conjuntos de nivel de $\mathbf{F}$ en U son las imágenes bajo $\mathbf{H}$ de los conjuntos de nivel planos de Φ en V. Son por tanto "superficies de dimensión-$(n - r)$" en U.

DETERMINANTES

Los determinantes son números que se asocian a las matrices cuadradas, y por lo tanto, a los operadores que están representados por dichas matrices. Son 0 si y solo si, el operador correspondiente no es invertible. Por lo tanto pueden usarse para decidir si las hipótesis de algunos de los teoremas anteriores se satisfacen. Tendrán un papel aún más importantes en el capítulo 10.

9.33 Definición Si $(j_1, \ldots, j_n)$ es una n-ada ordenada de enteros, definiremos

(82) $$s(j_1, \ldots, j_n) = \prod_{p<q} \operatorname{sgn}(j_q - j_p),$$

donde sgn $x = 1$ si $x > 0$, sgn $x = -1$ si $x < 0$, sgn $x = 0$ si $x = 0$. Entonces, $s(j_1, \ldots, j_n) = 1, -1$, o 0, y cambia de signo si se permuta algún par de j.

Sea $[A]$ la matriz de un operador lineal A en R^n relativa a la base estándar $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ con términos $a(i,j)$ en la fila i y columna j. Se define el determinante de $[A]$ como el número

(83) $$\det [A] = \sum s(j_1, \ldots, j_n) a(1, j_1) a(2, j_2) \cdots a(n, j_n).$$

La suma de (83) se extiende sobre todas las n-adas de enteros ordenadas $(j_1, \ldots, j_n)$ con $1 \le j_r \le n$.

Los vectores columna $\mathbf{x}_j$ de $[A]$ son

(84) $$\mathbf{x}_j = \sum_{i=1}^{n} a(i, j)\mathbf{e}_i \qquad (1 \le j \le n).$$

Será conveniente considerar a det $[A]$ como una función de los vectores columna de $[A]$. Si escribimos:

$$\det(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \det [A],$$

det es ahora una función real en el conjunto de todas las n-adas ordenadas de vectores en R^n.

9.34 Teorema

(*a*) *Si I es el operador identidad en R^n, será*

$$\det [I] = \det(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1.$$

(*b*) *det es una función lineal de cada uno de los vectores columna $\mathbf{x}_j$, si se mantienen los otros fijos.*

(*c*) *Si $[A]_1$ se obtiene de $[A]$ intercambiando dos columnas, entonces,* $\det [A]_1 = -\det [A]$.

(*d*) *Si $[A]$ tiene dos columnas iguales,* $\det [A] = 0$.

Demostración Si $A = I$, entonces $a(i,i) = 1$ y $a(i,j) = 0$ para $i \neq j$. Por tanto,

$$\det [I] = s(1, 2, \ldots, n) = 1,$$

lo que prueba (*a*). Por (82), $s(j_1, \ldots, j_n) = 0$ si dos cualesquiera de las j son iguales. Cada uno de los $n!$ productos restantes de (83) contiene un factor de cada columna, lo que prueba (*b*). El apartado (*c*)

es una consecuencia inmediata del hecho de cambiar de signo $s(j_1, \ldots, j_n)$ si se permutan dos j cualesquiera, y (*d*) es un corolario de (*c*).

9.35 Teorema *Si* $[A]$ *y* $[B]$ *son matrices n por n, entonces*

$$\det([B][A]) = \det[B] \det[A].$$

Demostración Si $\mathbf{x}_1, \ldots, \mathbf{x}_n$ son las columnas de $[A]$, definamos

(85) $$\Delta_B(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \Delta_B[A] = \det([B][A]).$$

Las columnas de $[B]\,[A]$ son los vectores $B\mathbf{x}_1, \ldots, B\mathbf{x}_n$. Así,

(86) $$\Delta_B(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \det(B\mathbf{x}_1, \ldots, B\mathbf{x}_n).$$

Por (86) y el Teorema 9.34, Δ_B tiene también las propiedades 9.34 (*b*) a (*d*). Por (*b*) y (84)

$$\Delta_B[A] = \Delta_B\left(\sum_i a(i, 1)\mathbf{e}_i, \mathbf{x}_2, \ldots, \mathbf{x}_n\right) = \sum_i a(i, 1)\,\Delta_B(\mathbf{e}_i, \mathbf{x}_2, \ldots, \mathbf{x}_n).$$

Repitiendo este proceso con $\mathbf{x}_2, \ldots, \mathbf{x}_n$, obtenemos

(87) $$\Delta_B[A] = \sum a(i_1, 1)a(i_2, 2) \cdots a(i_n, n)\,\Delta_B(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}),$$

estando extendida la suma sobre todas las n-tuplas ordenadas $(i_1, \ldots, i_n)$ con $1 \leq i_r \leq n$. Por (*c*) y (*d*),

(88) $$\Delta_B(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}) = t(i_1, \ldots, i_n)\,\Delta_B(\mathbf{e}_1, \ldots, \mathbf{e}_n),$$

donde $t = 1, 0,$ o -1, y como $[B][I] = [B]$, (57) muestra que

(89) $$\Delta_B(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \det[B].$$

Sustituyendo (89) y (88) en (87), obtenemos

$$\det([B][A]) = \left\{\sum a(i_1, 1) \cdots a(i_n, n)t(i_1, \ldots, i_n)\right\} \det[B],$$

para todas las matrices n por n $[A]$ y $[B]$. Tomando $B = I$, vemos que la suma anterior entre llaves es $\det[A]$, lo que prueba el teorema.

9.36 Teorema *Un operador lineal* A *en* R^n *es invertible si y solo si* $\det[A] \neq 0$.

Demostración Si A es invertible, el Teorema 9.35 muestra que

$$\det [A] \det [A^{-1}] = \det [AA^{-1}] = \det [I] = 1,$$

de modo que $\det [A] \neq 0$.

Si A no es invertible, las columnas $\mathbf{x}_1, \ldots, \mathbf{x}_n$ de $[A]$ son dependientes (Teorema 9.5); por tanto, hay uno, por ejemplo $\mathbf{x}_k$ tal que

(90) $$\mathbf{x}_k + \sum_{j \neq k} c_j \mathbf{x}_j = 0$$

para ciertos escalares c_j. Por 9.34 (*b*) y (*d*), se puede sustituir $\mathbf{x}_k$ por $\mathbf{x}_k + c_j \mathbf{x}_j$ sin alterar el determinante, si $j \neq k$. Repitiendo, vemos que puede sustituirse $\mathbf{x}_k$ por el primer miembro de (90), esto es, por $\mathbf{0}$, sin alterar el determinante. Pero una matriz que tiene una columna $\mathbf{0}$, tiene determinante 0. Por tanto, $\det [A] = 0$.

9.37 Observación Supongamos que $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ y $\{\mathbf{u}, \ldots, \mathbf{u}_n\}$ son bases en R^n. Todo operador lineal A en R^n determina matrices $[A]$ y $[A]u$, con términos a_{ij} y a_{ij} dados por

$$A\mathbf{e}_j = \sum_i a_{ij} \mathbf{e}_i, \qquad A\mathbf{u}_j = \sum_i \alpha_{ij} \mathbf{u}_i.$$

Si $\mathbf{u}_j = B\mathbf{e}_j = \Sigma b_{ij} \mathbf{e}_i$, entonces $A\mathbf{u}_j$ es igual a

$$\sum_k \alpha_{kj} B\mathbf{e}_k = \sum_k \alpha_{kj} \sum_i b_{ik} \mathbf{e}_i = \sum_i \left(\sum_k b_{ik} \alpha_{kj}\right) \mathbf{e}_i,$$

y también a

$$AB\mathbf{e}_j = A \sum_k b_{kj} \mathbf{e}_k = \sum_i \left(\sum_k a_{ik} b_{kj}\right) \mathbf{e}_i.$$

Así pues, $\Sigma b_{ik} \alpha_{kj} = \Sigma a_{ik} b_{kj}$, o

(91) $$[B][A]_U = [A][B].$$

Como B es invertible, $\det [B] \neq 0$. Por tanto, (91) juntamente con el Teorema 9.35 demuestra que

(92) $$\det [A]_U = \det [A].$$

El determinante de la matriz de un operador lineal no depende, por tanto, de la base que se utiliza para construir la matriz. *Es, pues, correcto hablar del determinante de un operador lineal, sin especificar ninguna base.*

9.38 Jacobianos Si $\mathbf{f}$ aplica un conjunto abierto $E \subset R^n$ en R^n, y es diferenciable en un punto $\mathbf{x} \in E$, el determinante del operador lineal $\mathbf{f}'(\mathbf{x})$ se le llama *Jacobiano de* $\mathbf{f}$ *en* $\mathbf{x}$. Simbólicamente:

(93) $$J_{\mathbf{f}}(\mathbf{x}) = \det \mathbf{f}'(\mathbf{x}).$$

También utilizaremos la expresión

$$\frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)} \tag{94}$$

para $J_f(\mathbf{x})$ si $(y_1, \ldots, y_n) = \mathbf{f}(x_1, \ldots, x_n)$.

En lenguaje jacobiano, la hipótesis principal del teorema de la función inversa es que $J_f(\mathbf{a}) \neq 0$ (compárese con el Teorema 9.36). Si la función implícita del teorema está establecida en términos de la función es (59), la hipótesis hecha para A equivale a

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} \neq 0.$$

DERIVADAS DE ORDEN SUPERIOR

9.39 Definición Supóngase que f es una función real definida en un conjunto abierto $E \subset R^n$, con derivadas parciales $D_1 f, \ldots, D_n f$. Si las funciones $D_j f$ son diferenciables, entonces las *derivadas parciales* de segundo orden de f están definidas como

$$D_{ij}f = D_i D_j f \qquad (i, j = 1, \ldots, n).$$

Si todas estas funciones $D_{ij}f$ son continuas en E, se dice que f es de clase $\mathscr{C}''$ en E, o que $f \in \mathscr{C}''(E)$.

Un mapeo $\mathbf{f}$ de E en R^m se dice que es de clase $\mathscr{C}''$ si cada componente de $\mathbf{f}$ es de clase $\mathscr{C}''$.

Puede suceder que en algún punto $D_{ij}f \neq D_{ji}f$, aunque ambas derivadas existan (véase Ejercicio 27). Sin embargo, se verá más adelante que $D_{ij}f = D_{ji}f$ siempre y cuando estas derivadas sean continuas.

Por simplicidad (y sin perder generalidad) se establecerán los dos teoremas siguientes para funciones reales de dos variables. El primero es un teorema de valor medio.

9.40 Teorema *Supóngase que f está definida en un conjunto abierto $E \subset R^2$, y que $D_1 f$ y $D_{21} f$ existen en cada punto de E. Supóngase también que $Q \subset E$ es un rectángulo cerrado con sus lados paralelos a los ejes coordenados, y que tienen a (a,b) y $(a + h, b + k)$ como vértices opuestos $(h \neq 0, k \neq 0)$. Si se hace*

$$\Delta(f, Q) = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b).$$

Entonces hay un punto (x,y) *en el interior de* Q *tal que*

(95) $$\Delta(f, Q) = hk(D_{21}f)(x, y).$$

Nótese la analogía entre (95) y el Teorema 5.10; el área de Q es hk.

Demostración Póngase $u(t) = f(t,b + k) - f(t,b)$. Dos aplicaciones del Teorema 5.10 muestran que hay un x entre a y $a + h$, y un y entre b y $b + k$, tales que

$$\begin{aligned}\Delta(f, Q) &= u(a + h) - u(a)\\ &= hu'(x)\\ &= h[(D_1f)(x, b + k) - (D_1f)(x, b)]\\ &= hk(D_{21}f)(x, y).\end{aligned}$$

9.41 Teorema *Supóngase que* f *está definida en un conjunto abierto* $E \subset R^2$, *y que* D_1f, $D_{21}f$ *y* D_2f *existen en cada punto de* E, *y* $D_{21}f$ *es continua en algún punto* $(a,b) \in E$.

Entonces D_{12} *existe en* (a,b) *y*

(96) $$(D_{12}f)(a, b) = (D_{21}f)(a, b).$$

Corolario $D_{21}f = D_{12}f$ *si* $f \in \mathscr{C}''(E)$.

Demostración Póngase $A = (D_{21}f)(a,b)$. Elíjase $\varepsilon > 0$. Si Q es el rectángulo del Teorema 9.40, y si h y k son suficientemente pequeños, se tiene que

$$|A - (D_{21}f)(x, y)| < \varepsilon$$

para todo $(x,y) \in Q$. Entonces, por (95)

$$\left|\frac{\Delta(f, Q)}{hk} - A\right| < \varepsilon,$$

Fíjese h y hágase $k \to 0$. Como D_2f existe en E, la última desigualdad implica que

(97) $$\left|\frac{(D_2f)(a + h, b) - (D_2f)(a, b)}{h} - A\right| \leq \varepsilon.$$

Como ε fue arbitraria, y debido a que (97) se cumple para todo $h \neq 0$ suficientemente pequeño, se deduce que $(D_{12}f)(a,b) = A$. Esto da (96).

DIFERENCIACIÓN DE INTEGRALES

Supóngase que φ es una función de dos variables que puede integrarse con respecto a una y diferenciarse con respecto a la otra. ¿Bajo qué condiciones se producirá el mismo resultado, si estos dos procesos de límite se llevan a cabo en orden opuesto? Para plantear la pregunta con más precisión: ¿Bajo qué condiciones sobre φ se puede demostrar que la ecuación

$$\frac{d}{dt}\int_a^b \varphi(x,t)\,dx = \int_a^b \frac{\partial\varphi}{\partial t}(x,t)\,dx \tag{98}$$

es verdadera? (El Ejercicio 28 proporciona un contra ejemplo.)

Será conveniente usar la notación

$$\varphi^t(x) = \varphi(x,t). \tag{99}$$

Entonces φ^t es para cada t, una función de una variable.

9.42 Teorema *Supóngase que*

(*a*) $\varphi(x,t)$ *está definida para* $a \le x \le b$, $c \le t \le d$;
(*b*) α *es una función creciente sobre* $[a,b]$;
(*c*) $\varphi^t \in \mathscr{R}(\alpha)$ *para cada* $t \in [c,d]$;
(*d*) $c < s < d$, *y para cada* $\varepsilon > 0$ *corresponde un* $\delta > 0$ *tal que*

$$|(D_2\varphi)(x,t) - (D_2\varphi)(x,s)| < \varepsilon$$

para todo $x \in [a,b]$ *y todo* $t \in (s-\delta, s+\delta)$.

Se define

$$f(t) = \int_a^b \varphi(x,t)\,d\alpha(x) \qquad (c \le t \le d). \tag{100}$$

Entonces $(D_2\varphi)^s \in \mathscr{R}(\alpha)$, $f'(s)$ *existe, y*

$$f'(s) = \int_a^b (D_2\varphi)(x,s)\,d\alpha(x). \tag{101}$$

Nótese que (*c*) solamente asegura la existencia de las integrales (100) para todo $t \in [c,d]$. Nótese también que (*d*) se cumple en realidad siempre que $D_2\varphi$ es continua sobre el rectángulo en el que φ está definida.

Demostración Consideremos los siguientes cocientes de diferencias:

$$\psi(x,t) = \frac{\varphi(x,t) - \varphi(x,s)}{t-s}$$

para $0 < |t - s| < \delta$. Por el Teorema 5.10, a cada (x,t) le corresponde un número u entre s y t tal que

$$\psi(x, t) = (D_2 \varphi)(x, u).$$

En consecuencia (*d*) implica que

(102) $$|\psi(x, t) - (D_2 \varphi)(x, s)| < \varepsilon \qquad (a \leq x \leq b, \quad 0 < |t - s| < \delta).$$

Nótese que

(103) $$\frac{f(t) - f(s)}{t - s} = \int_a^b \psi(x, t) \, d\alpha(x).$$

Y por (102), $\psi^t \to (D_2\varphi)^s$ uniformemente sobre $[a,b]$ cuando $t \to s$. Como cada $\psi^t \in \mathscr{R}(\alpha)$, entonces la conclusión deseada se deduce a partir de (103) y del Teorema 7.16.

9.43 Ejemplo Por supuesto que pueden demostrarse teoremas análogos al 9.42 con $(-\infty, \infty)$ en lugar de $[a,b]$. En lugar de hacer esto, veamos solamente un ejemplo. Se define

(104) $$f(t) = \int_{-\infty}^{\infty} e^{-x^2} \cos (xt) \, dx$$

y para $-\infty < t < \infty$,

(105) $$g(t) = -\int_{-\infty}^{\infty} x e^{-x^2} \operatorname{sen} (xt) \, dx,$$

Ambas integrales existen (convergen absolutamente) porque los valores absolutos de los integrandos son a lo más $\exp(-x^2)$ y $|x| \exp(-x^2)$, respectivamente.

Nótese que g se obtiene de f diferenciando el integrando con respecto a t. Se requiere que f sea diferenciable y que

(106) $$f'(t) = g(t) \qquad (-\infty < t < \infty).$$

Para demostrar esto, se examinarán primero los cocientes de diferencias del coseno: si $\beta > 0$ entonces

(107) $$\frac{\cos (\alpha + \beta) - \cos \alpha}{\beta} + \operatorname{sen} \alpha = \frac{1}{\beta} \int_{\alpha}^{\alpha + \beta} (\operatorname{sen} \alpha - \operatorname{sen} t) \, dt.$$

Como $|\operatorname{sen} \alpha - \operatorname{sen} t| \leq |t - \alpha|$, el miembro derecho de (107) es a lo más $\beta/2$ un valor absoluto; el caso $\beta < 0$ se trata de forma similar. Por esto

$$\left|\frac{\cos(\alpha+\beta)-\cos\alpha}{\beta}+\operatorname{sen}\alpha\right|\le|\beta| \tag{108}$$

para todo β (si el miembro izquierdo se interpreta como 0 cuando $\beta = 0$).

Ahora se fija t, y $h \neq 0$. Aplicando (108) con $\alpha = xt$, $\beta = xh$; se deduce entonces de (104) y (105) que

$$\left|\frac{f(t+h)-f(t)}{h}-g(t)\right|\le|h|\int_{-\infty}^{\infty}x^2e^{-x^2}\,dx.$$

Se obtiene entonces (106) cuando $h \to 0$.

Dando un paso más: la integración por partes aplicada a (104), muestra que

$$f(t)=2\int_{-\infty}^{\infty}xe^{-x^2}\frac{\operatorname{sen}(xt)}{t}\,dx. \tag{109}$$

Entonces $tf(t) = -2g(t)$, y (106) implica ahora que f satisface la ecuación diferencial

$$2f'(t)+tf(t)=0. \tag{110}$$

Si se resuelve esta ecuación diferencial y se considera que $f(0) = \sqrt{\pi}$ (véase Sec. 8.21), se encuentra que

$$f(t)=\sqrt{\pi}\exp\left(-\frac{t^2}{4}\right). \tag{111}$$

La integral (104) se determina entonces explícitamente.

EJERCICIOS

1. Si S es un subconjunto no vacío de un espacio vectorial X, demostrar (como se aseguró en la Sec. 9.1) que el generador de S es un espacio vectorial.
2. Demostrar (como se aseguró en la Sec. 9.6) que BA es lineal si A y B son transformaciones lineales.

   Demostrar también que A^{-1} es lineal e invertible.
3. Suponer que $A \in L(X, Y)$ y $A\mathbf{x} = \mathbf{0}$ solo cuando $\mathbf{x} = \mathbf{0}$. Demostrar que entonces A es 1-1.
4. Demostrar (como se aseguró en la Sec. 9.30) que los espacios nulos y los "rangos" de transformaciones lineales son espacios vectoriales.
5. Demostrar que a cada $A \in L(R^n, R^1)$ le corresponde un único $\mathbf{y} \in R^n$ tal que $A\mathbf{x} = \mathbf{x} \cdot \mathbf{y}$. Demostrar también que $\|A\| = |\mathbf{y}|$

Sugerencia: Bajo ciertas condiciones, en la desigualdad de Schwarz se cumple la igualdad.

6. Si $f(0, 0) = 0$, y

$$f(x, y) = \frac{xy}{x^2 + y^2} \qquad \text{si } (x, y) \neq (0, 0),$$

demostrar que $(D_1 f)(x,y)$ y $(D_2 f)(x,y)$ existen en cada punto de R^2, aunque f no sea continua en $(0, 0)$.

7. Suponer que f es una función de valores reales definida en un conjunto abierto $E \subset R^n$, y que las derivadas parciales $D_1 f, \ldots, D_n f$ están acotadas en E. Demostrar que f es continua en E.

Sugerencia: Procédase como en la demostración del Teorema 9.21.

8. Supóngase que f es una función real diferenciable en un conjunto abierto $E \subset R^n$, y que f tiene un máximo local en un punto $\mathbf{x} \in E$. Demostrar que $f'(\mathbf{x}) = 0$.

9. Si $\mathbf{f}$ es un mapeo diferenciable de un conjunto abierto *conexo* $E \subset R^n$ en R^m, y si $\mathbf{f}'(\mathbf{x}) = 0$ para cada $\mathbf{x} \in E$, demostrar que $\mathbf{f}$ es constante en E.

10. Si f es una función real definida en un conjunto abierto convexo $E \subset R^n$, tal que $(D_1 f)(\mathbf{x}) = 0$ para cada $\mathbf{x} \in E$, demostrar que $f(\mathbf{x})$ depende sólo de $x_2, \ldots, x_n$.

Mostrar que la convexibilidad de E puede reemplazarse por una condición más débil, pero que se requiere alguna condición. Por ejemplo, si $n = 2$ y E tiene la forma de una herradura, la proposición puede ser falsa.

11. Si f y g son funciones reales diferenciables en R^n, demostrar que

$$\nabla(fg) = f\nabla g + g\nabla f$$

y que $\nabla(1/f) = -f^{-2}\nabla f$, siempre y cuando $f \neq 0$.

12. Con dos números reales a y b fijos, $0 < a < b$. Defínase un mapeo $\mathbf{f} = (f_1, f_2, f_3)$ de R^2 en R^3 por medio de

$$f_1(s, t) = (b + a\cos s)\cos t$$
$$f_2(s, t) = (b + a\cos s)\,\text{sen}\; t$$
$$f_3(s, t) = a\,\text{sen}\, s.$$

Describir el "rango" K de $\mathbf{f}$. (Este es un subconjunto compacto de R^3.)

(*a*) Mostrar que hay exactamente 4 puntos $\mathbf{p} \in K$ tales que

$$(\nabla f_1)(\mathbf{f}^{-1}(\mathbf{p})) = \mathbf{0}.$$

Encontrar estos puntos.

(*b*) Determinar el conjunto de todos los $\mathbf{q} \in K$ tales que

$$(\nabla f_3)(\mathbf{f}^{-1}(\mathbf{q})) = \mathbf{0}.$$

(*c*) Mostrar que uno de los puntos $\mathbf{p}$ encontrados en la parte (*a*) corresponde a un mínimo local de f_1, uno a un máximo local y los otros dos restantes a nada (estos se llaman "puntos silla").

¿Cuáles de los puntos $\mathbf{q}$ determinados en la parte (*b*) corresponden a máximos o mínimos?

(*d*) Sea λ un número real irracional, y defínase $\mathbf{g}(t) = f(t), \lambda t)$. Demostrar que $\mathbf{g}$ es un mapeo 1-1 de R^1 sobre un subconjunto denso de K. Demostrar que

$$|\mathbf{g}'(t)|^2 = a^2 + \lambda^2(b + a\cos t)^2.$$

13. Supóngase que $\mathbf{f}$ es un mapeo diferenciable de R^1 en R^3 tal que $|\mathbf{f}(t)| = 1$ para cada t. Demostrar que $\mathbf{f}'(t) \cdot \mathbf{f}(t) = 0$.

Interpretar este resultado geométricamente.

14. Definir $f(0, 0) = 0$, y

$$f(x, y) = \frac{x^3}{x^2 + y^2} \qquad \text{si } (x, y) \neq (0, 0).$$

(*a*) Demostrar que $D_1 f$ y $D_2 f$ son funciones acotadas en R^2. (Por consiguiente f es continua.)

(*b*) Sea $\mathbf{u}$ un vector unitario en R^2. Mostrar que la derivada direccional $(D_{\mathbf{u}}f)(0, 0)$ existe, y que a lo más su valor absoluto es 1.

(*c*) Sea γ un mapeo diferenciable de R^1 en R^2 (en otras palabras, γ es una curva diferenciable en R^2), con $\gamma(0) = (0, 0)$ y $|\gamma'(0)| > 0$. Si se pone $g(t) = f(\gamma(t))$, demostrar que g es diferenciable para cada $t \in R^1$.

Si $\gamma \in \mathscr{C}'$ demostrar que $g \in \mathscr{C}'$.

(*d*) A pesar de esto, demostrar que f no es diferenciable en $(0, 0)$.

Sugerencia: La fórmula (40) falla.

15. Si se define $f(0, 0) = 0$, y se hace

$$f(x, y) = x^2 + y^2 - 2x^2 y - \frac{4x^6 y^2}{(x^4 + y^2)^2}$$

para $(x,y) \neq (0, 0)$.

(*a*) Demostrar, para todo $(x,y) \in R^2$, que

$$4x^4 y^2 \leq (x^4 + y^2)^2.$$

Concluir de esto que f es continua.

(*b*) Para $0 \leq \theta \leq 2\pi$, $-\infty < t < \infty$, se define

$$g_\theta(t) = f(t\cos\theta, t \operatorname{sen}\theta).$$

Mostrar que $g_\theta(0) = 0$, $g_\theta'(0) = 0$, $g_\theta''(0) = 2$. Cada g_θ tiene por lo tanto, un mínimo local estricto en $t = 0$.

En otras palabras, la restricción de f a cada recta que pasa por $(0, 0)$ tiene un mínimo local estricto en $(0, 0)$.

(*c*) Mostrar que $(0, 0)$ no es sin embargo un mínimo local para f, debido a que $f(x, x^2) = -x^4$.

16. Mostrar que la continuidad de $\mathbf{f}'$ en el punto $\mathbf{a}$ es necesaria en el teorema de la función inversa, inclusive en el caso $n = 1$: si

$$f(t) = t + 2t^2 \operatorname{sen}\left(\frac{1}{t}\right)$$

para $t \neq 0$, y $f(0) = 0$, entonces $f'(0) = 1$, f' es acotada en $(-1, 1)$, pero f no es uno-a-uno en cualquier vecindad de 0.

17. Sea $\mathbf{f} = (f_1, f_2)$ un mapeo de R^2 en R^2 dado por

$$f_1(x, y) = e^x \cos y, \qquad f_2(x, y) = e^x \operatorname{sen} y.$$

(*a*) ¿Cuál es el "rango" de f?

(*b*) Mostrar que el Jacobiano de f no es 0 en ningún punto de R^2. Entonces cada punto de R^2 tiene una vecindad en la que f es uno-a-uno, no obstante que f no es uno-a-uno sobre R^2.

(*c*) Poniendo $\mathbf{a} = (0, \pi/3)$, $\mathbf{b} = f(\mathbf{a})$, sea $\mathbf{g}$ el inverso continuo de $\mathbf{f}$, definido en una vecindad de $\mathbf{b}$, tal que $\mathbf{g}(\mathbf{b}) = \mathbf{a}$. Determinar una fórmula explícita para $\mathbf{g}$, calcular también $\mathbf{f}'(\mathbf{a})$ y $\mathbf{g}'(\mathbf{b})$, y verificar la fórmula (52).

(*d*) ¿Cuáles son las imágenes bajo $\mathbf{f}$ de las rectas paralelas a los ejes coordenados?

18. Conteste las preguntas análogas para el mapeo definido por

$$u = x^2 - y^2, \qquad v = 2xy.$$

19. Mostrar que el sistema de ecuaciones

$$\begin{aligned} 3x + y - z + u^2 &= 0 \\ x - y + 2z + u &= 0 \\ 2x + 2y - 3z + 2u &= 0 \end{aligned}$$

puede resolverse para x, y, u en términos de z; para x, z, u en términos de y; para y, z, u en términos de x; pero no para x, y, z en términos de u.

20. Hacer $n = m = 1$ en el teorema de la función implícita e interpretar gráficamente el teorema (y también su demostración).

21. Sea f definida en R^2 por

$$f(x, y) = 2x^3 - 3x^2 + 2y^3 + 3y^2.$$

(*a*) Encontrar los cuatro puntos de R^2 en los cuales el gradiente de f es cero. Mostrar que f tiene exactamente un máximo y un mínimo locales en R^2.

(*b*) Si S es el conjunto de todos los $(x,y) \in R^2$ en los cuales $f(x,y) = 0$. Encontrar los puntos de S que no tienen vecindades para las cuales la ecuación $f(x,y) = 0$ puede resolverse para y en términos de x (o para x en términos de y). Describir S de la forma más precisa que se pueda.

22. Hacer algo similar para

$$f(x, y) = 2x^3 + 6xy^2 - 3x^2 + 3y^2.$$

23. Si se define f en R^3 por medio de

$$f(x, y_1, y_2) = x^2 y_1 + e^x + y_2.$$

Mostrar que $f(0, 1, -1) = 0$, $(D_1 f)(0, 1, -1) \neq 0$, y que existe por lo tanto una función diferenciable g en alguna vecindad de $(1, -1)$ en R^2, tal que $g(1, -1) = 0$ y

$$f(g(y_1, y_2), y_1, y_2) = 0.$$

Encontrar $(D_1 g)(1, -1)$ y $(D_2 g)(1, -1)$.

24. Para $(x,y) \neq (0, 0)$, se define $\mathbf{f} = (f_1, f_2)$ como

$$f_1(x, y) = \frac{x^2 - y^2}{x^2 + y^2}, \qquad f_2(x, y) = \frac{xy}{x^2 + y^2}.$$

Calcular el rango de $\mathbf{f}'(x,y)$, y encontrar el "rango" de $\mathbf{f}$.

25. Supóngase que $A \in L(R^n, R^m)$, y sea r el rango de A.

(a) Si se define S como en la demostración del Teorema 9.32, mostrar que SA es una proyección en R^n cuyo espacio nulo es $\mathscr{N}(A)$ y cuyo "rango" es $\mathscr{R}(S)$. *Sugerencia*: Por (68), $SASA = SA$.

(b) Usar (a) para mostrar que

$$\dim \mathscr{N}(A) + \dim \mathscr{R}(A) = n.$$

26. Mostrar que la existencia (e incluso la continuidad) de $D_{12}f$ no implica la existencia de $D_1 f$. Por ejemplo, sea $f(x,y) = g(x)$, donde g es en ninguna parte diferenciable.

27. Si se hace $f(0, 0) = 0$, y

$$f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}$$

para $(x,y) \neq (0, 0)$. Demostrar que

(a) f, $D_1 f$, $D_2 f$ son continuas en R^2;

(b) $D_{12}f$ y $D_{21}f$ existen en cada punto de R^2, y son continuas excepto en $(0, 0)$;

(c) $(D_{12}f)(0, 0) = 1$, y $(D_{21}f)(0, 0) = -1$.

28. Para $t \geq 0$, se hace

$$\varphi(x, t) = \begin{cases} x & (0 \leq x \leq \sqrt{t}) \\ -x + 2\sqrt{t} & (\sqrt{t} \leq x \leq 2\sqrt{t}) \\ 0 & (\text{en otro caso}), \end{cases}$$

y se pone $\varphi(x,t) = -\varphi(x, |t|)$ si $t < 0$.

Mostrar que φ es continua sobre R^2, y que

$$(D_2 \varphi)(x, 0) = 0$$

para todo x. Ahora defínase

$$f(t) = \int_{-1}^{1} \varphi(x, t)\, dx.$$

Mostrar que $f(t) = t$ si $|t| < \frac{1}{4}$. En consecuencia,

$$f'(0) \neq \int_{-1}^{1} (D_2 \varphi)(x, 0)\, dx.$$

29. Sea E un conjunto abierto en R^n. Las clases $\mathscr{C}'(E)$ y $\mathscr{C}''(E)$ se han definido en el texto. Por inducción, puede definirse $\mathscr{C}^{(k)}(E)$ de la forma siguiente, para todos los enteros positivos k: Decir que $f \in \mathscr{C}^{(k)}(E)$ significa que las derivadas parciales $D_1 f, \ldots, D_n f$ pertenecen a $\mathscr{C}^{(k-1)}(E)$.

Admitir que $f \in \mathscr{C}^{(k)}(E)$, y mostrar (aplicando repetidamente el Teorema 9.41) que la derivada de orden k-ésimo

$$D_{i_1 i_2 \ldots i_k} f = D_{i_1} D_{i_2} \ldots D_{i_k} f$$

no cambia si los subíndices $i_1, \ldots, i_k$ se permutan.

Por ejemplo, si $n \geq 3$, entonces

$$D_{1213} f = D_{3112} f$$

para cada $f \in \mathscr{C}^{(4)}$.

30. Sea $f \in \mathscr{C}^{(m)}(E)$, donde E es un subconjunto abierto de R^n. Fíjese $\mathbf{a} \in E$, y supóngase que $\mathbf{x} \in R^n$ es tan cercano a $\mathbf{0}$ que los puntos

$$\mathbf{p}(t) = \mathbf{a} + t\mathbf{x}$$

están en E siempre que $0 \leq t \leq 1$. Si se define

$$h(t) = f(\mathbf{p}(t))$$

para todo $t \in R^1$ para los que $\mathbf{p}(t) \in E$.

(a) Mostrar que para $1 \leq k \leq m$, (aplicando repetidamente la regla de la cadena)

$$h^{(k)}(t) = \sum (D_{i_1 \ldots i_k} f)(\mathbf{p}(t))\, x_{i_1} \ldots x_{i_k}.$$

La suma se extiende sobre todas las k-adas ordenadas $(i_1, \ldots, i_k)$ en las que i_j es uno de los enteros $1, \ldots, n$.

(b) Del teorema de Taylor (5.15),

$$h(1) = \sum_{k=0}^{m-1} \frac{h^{(k)}(0)}{k!} + \frac{h^{(m)}(t)}{m!}$$

para algún $t \in (0, 1)$. Usar esto para demostrar el teorema de Taylor en n variables, mostrando que la fórmula

$$f(\mathbf{a} + \mathbf{x}) = \sum_{k=0}^{m-1} \frac{1}{k!} \sum (D_{i_1 \ldots i_k} f)(\mathbf{a}) x_{i_1} \cdots x_{i_k} + r(\mathbf{x})$$

representa a $f(\mathbf{a} + \mathbf{x})$ como la suma del llamado "polinomio de Taylor de grado $m - 1$", más un residuo que satisface la condición

$$\lim_{\mathbf{x} \to \mathbf{0}} \frac{r(\mathbf{x})}{|\mathbf{x}|^{m-1}} = 0.$$

Cada una de las sumas internas se extiende sobre todas las k-adas $(i_1, \ldots, i_k)$ ordenadas, de la misma manera que en (a); como de costumbre, la derivada

de f de orden cero es simplemente f, así que el término constante del polinomio de Taylor de f en $\mathbf{a}$ es $f(\mathbf{a})$.

(c) El Ejercicio 29 muestra que la repetición ocurre en el polinomio de Taylor cuando se escribe como en (b). Por ejemplo, D_{113} aparece tres veces, como D_{113}, D_{131}, D_{311}. La suma de los tres términos correspondientes se puede escribir en la forma

$$3(D_1^2 D_3 f)(\mathbf{a})x_1^2 x_3 .$$

Demostrar (calculando la frecuencia en la que cada derivada aparece) que el polinomio de Taylor de (b) se puede escribir en la forma

$$\sum \frac{(D_1^{s_1} \cdots D_n^{s_n} f)(\mathbf{a})}{s_1! \cdots s_n!} x_1^{s_1} \cdots x_n^{s_n}.$$

Aquí la suma se extiende sobre todas las n-adas ordenadas $(s_1, \ldots, s_n)$ tales que cada s_i es un entero no negativo, y $s_1 + \cdots + s_n \leq m - 1$.

31. Supóngase que $f \in \mathscr{C}^{(3)}$ en alguna vecindad de un punto $\mathbf{a} \in R^2$, el gradiente de f es 0 en $\mathbf{a}$, pero que no todas las derivadas de f de segundo orden son 0 en $\mathbf{a}$. Muestre cómo pueden determinarse entonces a partir del polinomio de Taylor de f en $\mathbf{a}$ (de grado 2) si f tiene un máximo local, o un mínimo local, o ninguna de las dos cosas, en el punto $\mathbf{a}$.

Extender esto a R^n en lugar de R^2.

10

INTEGRACIÓN DE FORMAS DIFERENCIALES

La integración puede estudiarse en varios niveles. En el capítulo 6, se desarrolló la teoría para funciones que se comportan razonablemente bien sobre subintervalos de la recta real. En el capítulo 11 se encontrará una teoría de integración con un desarrollo bastante avanzado que se puede aplicar a clases de funciones más grandes, cuyos dominios son conjuntos más o menos arbitrarios, y no necesariamente subconjuntos de R^n. Este capítulo se dedica a los aspectos de la teoría de integración que están muy relacionados con la geometría de espacios euclidianos, como la fórmula de cambio de variable, integrales de línea y los mecanismos de las formas diferenciales que se usan en las proposiciones y demostración del teorema en n-dimensiones, análogo al teorema fundamental del cálculo, es decir el teorema de Stokes.

INTEGRACIÓN

10.1 Definición Supongamos que I^k es una k-celda en R^k que consta de todos los

$$\mathbf{x} = (x_1, \ldots, x_k)$$

tales que

$$a_i \le x_i \le b_i \qquad (i = 1, \ldots, k), \tag{1}$$

I^j es la j-celda en R^j definida por las j primeras desigualdades (1) y f es una función real continua en I^k.

Pongamos $f = f_k$, y definamos f_{k-1} en I^{k-1} por

$$f_{k-1}(x_1, \ldots, x_{k-1}) = \int_{a_k}^{b_k} f_k(x_1, \ldots, x_{k-1}, x_k)\, dx_k .$$

La continuidad uniforme de f_k en I^k demuestra que f_{k-1} es continua en I^{k-1}. Por tanto, podemos repetir este proceso y obtener funciones f_j continuas en I^j, tales que f_{j-1} es la integral de f_j con respecto a x_j, sobre $[a_j, b_j]$. Después de k pasos llegamos a un *número* f_0, que llamamos *integral de* f *sobre* I^k; lo escribiremos en la forma

(2) $$\int_{I^k} f(\mathbf{x})\, d\mathbf{x} \qquad \text{or} \qquad \int_{I^k} f.$$

A priori, esta definición de integral depende del orden en que se lleven a cabo las k integraciones. Sin embargo, esta dependencia es solo aparente. Para demostrarlo, introduzcamos la notación provisional $L(f)$ para la integral (2) y $L'(f)$ para el resultado obtenido llevando a cabo las k integraciones en algún otro orden.

10.2 Teorema *Para todo* $f \in \mathscr{C}(I^k)$, $L(f) = L'(f)$.

Demostración Si $h(\mathbf{x}) = h_1(x_1) \cdots h_k(x_k)$, donde $h_j \in \mathscr{C}([a_j, b_j])$, entonces

$$L(h) = \prod_{i=1}^{k} \int_{a_i}^{b_i} h_i(x_i)\, dx_i = L'(h).$$

Si $\mathscr{A}$ es el conjunto de todas las sumas finitas de tales funciones h, se deduce que $L(g) = L'(g)$ para todo $g \in \mathscr{A}$. Además, $\mathscr{A}$ es un álgebra de funciones en I^k para la cual es aplicable el teorema de Stone-Weierstrass.

Hagamos $V = \prod_1^k (b_i - a_i)$. Si $f \in \mathscr{C}(I^k)$ y $\varepsilon > 0$, existe $g \in \mathscr{A}$ tal que $\|f - g\| < \varepsilon/V$, donde $\|f\|$ está definida como igual a máx $|f(\mathbf{x})|$ $(\mathbf{x} \in I^k)$. Entonces $\|L(f - g)| < \varepsilon$, $|L'(f - g| < \varepsilon$, y como

$$L(f) - L'(f) = L(f - g) + L'(g - f),$$

deducimos que $|L(f) - L'(f)| < 2\varepsilon$.

En relación con esto, es importante el Ejercicio 2.

10.3 Definición El *soporte* de una función (real o compleja) f en R^k es la cerradura del conjunto de todos los puntos $\mathbf{x} \in R^k$ en los que $f(\mathbf{x}) \neq 0$. Si f

es una función continua con soporte compacto, sea I^k alguna k-celda que contiene al soporte de f, y definamos

(3)
$$\int_{R^k} f = \int_{I^k} f.$$

La integral así definida es, evidentemente, independiente de la elección de I^k, con tal, solamente, de que I^k contenga al soporte de f.

Es tentador, ahora, extender la definición de integral sobre R^k a las funciones que son límite (en algún sentido) de funciones continuas con soporte compacto. No necesitamos estudiar las condiciones bajo las cuales puede hacerse esto; el enunciado apropiado de esta cuestión es la integral de Lebesgue. Nos limitaremos a describir un ejemplo muy sencillo que se utilizará en la demostración del teorema de Stokes.

10.4 Ejemplo Sea Q^k el k-simplex que consta de todos los puntos $\mathbf{x} = (x_1, \ldots, x_k)$ en R^k para los cuales $x_1 + \cdots + x_k \leq 1$ y $x_i \geq 0$ para $i = 1, \ldots, k$. Si $k = 3$, por ejemplo, Q^k es un tetraedro, con vértices en $\mathbf{0}$, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Si $f \in \mathscr{C}(Q^k)$, extenderemos f a una función en I^k poniendo $f(\mathbf{x}) = 0$ fuera de Q^k y definiremos

(4)
$$\int_{Q^k} f = \int_{I^k} f.$$

Aquí, I^k es el «cubo unidad» definido por

$$0 \leq x_i \leq 1 \qquad (1 \leq i \leq k).$$

Como f puede ser discontinua en I^k, la existencia de la integral del segundo miembro de (4) necesita ser demostrada. También queremos mostrar que esta integral es independiente del orden en que se lleva a cabo cada una de las k integraciones.

Para hacer esto, supóngase que $0 < \delta < 1$, poniendo

(5)
$$\varphi(t) = \begin{cases} 1 & (t \leq 1 - \delta) \\ \dfrac{(1 - t)}{\delta} & (1 - \delta < t \leq 1) \\ 0 & (1 < t), \end{cases}$$

y se define

(6)
$$F(\mathbf{x}) = \varphi(x_1 + \cdots + x_k) f(\mathbf{x}) \qquad (\mathbf{x} \in I^k).$$

Entonces $F \in \mathscr{C}(I^k)$.

Se hace $\mathbf{y} = (x_1, \ldots, x_{k-1})$, $\mathbf{x} = (\mathbf{y}, x_k)$. Para cada $\mathbf{y} \in I^{k-1}$, el conjunto de todos los x_k tales que $F(\mathbf{y}, x_k) \neq f(\mathbf{y}; x_k)$ es ya sea vacío o un segmento cuya longitud no excede a la de δ. Como $0 \leq \varphi \leq 1$, se deduce que

$$(7) \qquad |F_{k-1}(\mathbf{y}) - f_{k-1}(\mathbf{y})| \leq \delta \|f\| \qquad (\mathbf{y} \in I^{k-1}),$$

donde $\|f\|$ tiene el mismo significado que en la demostración del Teorema 10.2, y F_{k-1}, f_{k-1} están dadas como en la Definición 10.1.

Cuando $\delta \to 0$, (7) exhibe a f_{k-1} como un límite uniforme de una sucesión de funciones continuas. Por esto $f_{k-1} \in \mathscr{C}(I^{k-1})$, y las integraciones adicionales no presentan problema.

Esto demuestra la existencia de la integral (4). Además, (7) muestra que

$$(8) \qquad \left| \int_{I^k} F(\mathbf{x})\, d\mathbf{x} - \int_{I^k} f(\mathbf{x})\, d\mathbf{x} \right| \leq \delta \|f\|.$$

Nótese que (8) es verdadero, no importando el orden en el cual se efectúen las k integraciones sencillas. Como $F \in \mathscr{C}(I^k)$, $\int F$ no es afectada por ningún cambio en este orden. En consecuencia (8) muestra que lo mismo es válido para $\int f$.

Esto completa la demostración.

Nuestro siguiente objetivo es la fórmula de cambio de variable que se establece en el Teorema 10.9. Para facilitar su demostración, se estudiarán primero los denominados mapeos primitivos, y las particiones de la unidad. Los mapeos primitivos permitirán obtener una visión clara de la acción local de un mapeo- $\mathscr{C}'$ con derivada invertible, y las particiones de la unidad son un recurso muy útil que hace posible usar la información local dentro de un marco global.

MAPEOS PRIMITIVOS

10.5 Definición Si $\mathbf{G}$ mapea un conjunto abierto $E \subset R^n$ en R^n, y si hay un entero m y una función real g con dominio E tal que

$$(9) \qquad \mathbf{G}(\mathbf{x}) = \sum_{i \neq m} x_i \mathbf{e}_i + g(\mathbf{x})\mathbf{e}_m \qquad (\mathbf{x} \in E),$$

entonces $\mathbf{G}$ se denomina *primitivo*. Un mapeo primitivo es entonces aquel que cambia a lo más una coordenada. Nótese que (9) puede escribirse también en la forma

$$(10) \qquad \mathbf{G}(\mathbf{x}) = \mathbf{x} + [g(\mathbf{x}) - x_m]\mathbf{e}_m.$$

Si g es diferenciable en algún punto $\mathbf{a} \in E$, también lo es $\mathbf{G}$. La matriz $[\alpha_{ij}]$ del operador $\mathbf{G}'(\mathbf{a})$ tiene como m-ésimo renglón a

(11) $$(D_1 g)(\mathbf{a}), \ldots, (D_m g)(\mathbf{a}), \ldots, (D_n g)(\mathbf{a})$$

para $j \neq m$, se tiene $\alpha_{jj} = 1$ y $\alpha_{ij} = 0$ si $i \neq j$. El Jacobiano de $\mathbf{G}$ en $\mathbf{a}$ está dado entonces por

(12) $$J_{\mathbf{G}}(\mathbf{a}) = \det[\mathbf{G}'(\mathbf{a})] = (D_m g)(\mathbf{a}),$$

y se ve (por el Teorema 9.36) que $\mathbf{G}'(\mathbf{a})$ *es invertible si y solo si* $(D_m g)(\mathbf{a}) \neq$ $\neq 0$.

10.6 Definición Un operador lineal B sobre R^n que intercambia algún par de miembros de la base estándar y deja fijos a otros, se denominá un *flip*.

Por ejemplo, el flip B sobre R^4 *que intercambia* $\mathbf{e}_2$ y $\mathbf{e}_4$ tiene la forma

(13) $$B(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 + x_4\mathbf{e}_4) = x_1\mathbf{e}_1 + x_2\mathbf{e}_4 + x_3\mathbf{e}_3 + x_4\mathbf{e}_2$$

o de manera equivalente,

(14) $$B(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 + x_4\mathbf{e}_4) = x_1\mathbf{e}_1 + x_4\mathbf{e}_2 + x_3\mathbf{e}_3 + x_2\mathbf{e}_4.$$

Por consiguiente B puede considerarse como un intercambio de dos de las coordenadas, en lugar de dos vectores base.

En la siguiente demostración se usarán las proyecciones $P_0, \ldots, P_n$ en R^n, definidas por $P_0\mathbf{x} = \mathbf{0}$ y

(15) $$P_m\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_m\mathbf{e}_m$$

para $1 \leq m \leq n$. Entonces P_m es la proyección cuyo rango y espacio nulo están generados por $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ y $\{\mathbf{e}_{m+1}, \ldots, \mathbf{e}_n\}$ respectivamente.

10.7 Teorema *Supóngase que* $\mathbf{F}$ *es un mapeo-*$\mathscr{C}'$ *de un conjunto abierto* $E \subset R^n$ *en* R^n, $\mathbf{0} \in E$, $\mathbf{F}(\mathbf{0}) = \mathbf{0}$, *y* $\mathbf{F}'(\mathbf{0})$ *es invertible.*

Entonces hay una vecindad de $\mathbf{0}$ *en* R^n *en la cual es válida una representación de la forma*

(16) $$\mathbf{F}(\mathbf{x}) = B_1 \cdots B_{n-1}\mathbf{G}_n \circ \cdots \circ \mathbf{G}_1(\mathbf{x})$$

En (16) *cada* $\mathbf{G}_i$ *es un mapeo-* $\mathscr{C}'$ *primitivo en alguna vecindad de* $\mathbf{0}$; $\mathbf{G}_i(\mathbf{0}) = \mathbf{0}$, $\mathbf{G}_i'(\mathbf{0})$ *es invertible, y cada* B_i *es ya sea un flip o un operador identidad.*

Abreviando, (16) representa a $\mathbf{F}$ localmente como composición de mapeos primitivos y flips.

Demostración Poniendo $\mathbf{F} = \mathbf{F}_1$. Admítase que $1 \le m \le n - 1$, y hágase la siguiente hipótesis de inducción (que evidentemente se cumple para $m = 1$):

V_m es una vecindad de $\mathbf{0}$, $\mathbf{F}_m \in \mathscr{C}'(V_m)$, $\mathbf{F}_m(\mathbf{0}) = \mathbf{0}$, $\mathbf{F}_m'(\mathbf{0})$ *es invertible, y*

(17) $$P_{m-1}\mathbf{F}_m(\mathbf{x}) = P_{m-1}\mathbf{x} \qquad (\mathbf{x} \in V_m).$$

De (17), se tiene

(18) $$\mathbf{F}_m(\mathbf{x}) = P_{m-1}\mathbf{x} + \sum_{i=m}^{n} \alpha_i(\mathbf{x})\mathbf{e}_i,$$

en donde $\alpha_m, \ldots, \alpha_n$ son funciones- $\mathscr{C}'$ reales en V_m. Por consiguiente,

(19) $$\mathbf{F}_m'(\mathbf{0})\mathbf{e}_m = \sum_{i=m}^{n} (D_m \alpha_i)(\mathbf{0})\mathbf{e}_i.$$

Como $\mathbf{F}_m'(\mathbf{0})$ es invertible, el miembro izquierdo de (19) no es $\mathbf{0}$, y por lo tanto hay un k tal que $m \le k \le n$ y $(D_m \alpha_k)(\mathbf{0}) \ne 0$.

Sea B_m el flip que intercambia m y este k (si $k = m$, B_m es la identidad) y defínase

(20) $$\mathbf{G}_m(\mathbf{x}) = \mathbf{x} + [\alpha_k(\mathbf{x}) - x_m]\mathbf{e}_m \qquad (\mathbf{x} \in V_m).$$

Entonces $\mathbf{G}_m \in \mathscr{C}'(V_m)$, $\mathbf{G}_m$ es primitivo, y $\mathbf{G}_m'(\mathbf{0})$ es invertible, ya que $(D_m \alpha_k)(\mathbf{0}) \ne 0$.

Por tanto, el teorema de la función inversa muestra que hay un conjunto abierto U_m, con $\mathbf{0} \in U_m \subset V_m$, tal que $\mathbf{G}_m$ es un mapeo 1-1 de U_m sobre una vecindad V_{m+1} de $\mathbf{0}$, en la cual $\mathbf{G}_m^{-1}$ es continuamente diferenciable. Si se define $\mathbf{F}_{m+1}$ como

(21) $$\mathbf{F}_{m+1}(\mathbf{y}) = B_m \mathbf{F}_m \circ \mathbf{G}_m^{-1}(\mathbf{y}) \qquad (\mathbf{y} \in V_{m+1}).$$

Entonces $\mathbf{F}_{m+1} \in \mathscr{C}'(V_{m+1})$, $\mathbf{F}_{m+1}(\mathbf{0}) = \mathbf{0}$, y $\mathbf{F}_{m+1}'(\mathbf{0})$ es invertible (debido a la regla de la cadena). También para $\mathbf{x} \in U_m$,

(22) $$\begin{aligned} P_m \mathbf{F}_{m+1}(\mathbf{G}_m(\mathbf{x})) &= P_m B_m \mathbf{F}_m(\mathbf{x}) \\ &= P_m[P_{m-1}\mathbf{x} + \alpha_k(\mathbf{x})\mathbf{e}_m + \cdots] \\ &= P_{m-1}\mathbf{x} + \alpha_k(\mathbf{x})\mathbf{e}_m \\ &= P_m \mathbf{G}_m(\mathbf{x}) \end{aligned}$$

así que

(23) $$P_m \mathbf{F}_{m+1}(\mathbf{y}) = P_m \mathbf{y} \qquad (\mathbf{y} \in V_{m+1}).$$

La hipótesis de inducción se cumple por tanto con $m + 1$ en lugar de m.

[En (22) se usó primero (21), después (18) y la definición de B_m; después, la definición de P_m, y finalmente (20).]

Debido a que $B_m B_m = I$, (21) con $\mathbf{y} = \mathbf{G}_m(\mathbf{x})$, es equivalente a

$$\mathbf{F}_m(\mathbf{x}) = B_m \mathbf{F}_{m+1}(\mathbf{G}_m(\mathbf{x})) \qquad (\mathbf{x} \in U_m). \tag{24}$$

Si se aplica esto con $m = 1, \ldots, n - 1$, se obtiene

$$\begin{aligned} \mathbf{F} = \mathbf{F}_1 &= B_1 \mathbf{F}_2 \circ \mathbf{G}_1 \\ &= B_1 B_2 \mathbf{F}_3 \circ \mathbf{G}_2 \circ \mathbf{G}_1 = \cdots \\ &= B_1 \cdots B_{n-1} \mathbf{F}_n \circ \mathbf{G}_{n-1} \circ \cdots \circ \mathbf{G}_1 \end{aligned}$$

en alguna vecindad de $\mathbf{0}$. Por (17), $\mathbf{F}_n$ es primitivo. Esto completa la demostración.

PARTICIONES DE LA UNIDAD

10.8 Teorema *Supóngase que K es un subconjunto compacto de R^n, y $\{V_\alpha\}$ es una cubierta abierta de K. Entonces existen funciones $\psi_1, \ldots, \psi_s \in$* $\mathscr{C}(R^n)$ tales que

(*a*) $0 \le \psi_i \le 1$ *para* $1 \le i \le s$;
(*b*) *cada ψ_i tiene su soporte en algún V_α, y*
(*c*) $\psi_1(\mathbf{x}) + \cdots + \psi_s(\mathbf{x}) = 1$ *para cada* $\mathbf{x} \in K$.

A causa de (*c*), $\{\psi_i\}$ se llama una *partición de la unidad*, y (*b*) se expresa algunas veces diciendo que $\{\psi_i\}$ está *subordinada* a la cubierta $\{V_\alpha\}$.

Corolario *Si $f \in \mathscr{C}(R^n)$ y el soporte de f está en K, entonces*

$$f = \sum_{i=1}^{s} \psi_i f. \tag{25}$$

Cada ψ_i tiene su soporte en algún V_α.

Lo importante de (25) es que proporciona una representación de f como una suma de funciones continuas $\psi_i f$ con soportes "pequeños".

Demostración Se asocia con cada $\mathbf{x} \in K$ un índice $\alpha(\mathbf{x})$ de manera que $\mathbf{x} \in V_{\alpha(\mathbf{x})}$. Entonces hay bolas abiertas $B(\mathbf{x})$ y $W(\mathbf{x})$, con centro en $\mathbf{x}$, tales que

$$\overline{B(\mathbf{x})} \subset W(\mathbf{x}) \subset \overline{W(\mathbf{x})} \subset V_{\alpha(\mathbf{x})}. \tag{26}$$

Como K es compacto, hay puntos $\mathbf{x}_1, \ldots, \mathbf{x}_s$ en K tales que

(27) $$K \subset B(\mathbf{x}_1) \cup \cdots \cup B(\mathbf{x}_s).$$

Por (26) hay funciones $\varphi_1, \ldots, \varphi_s \in \mathscr{C}(R^n)$, tales que $\varphi_i(\mathbf{x}) = 1$ sobre $B(\mathbf{x}_i)$, $\varphi_i(\mathbf{x}) = 0$ fuera de $W(x_i)$, y $0 \leq \varphi_i(\mathbf{x}) \leq 1$ sobre R^n. Se define $\psi_1 = \varphi_1$ y

(28) $$\psi_{i+1} = (1 - \varphi_1) \cdots (1 - \varphi_i)\varphi_{i+1}$$

para $i = 1, \ldots, s - 1$.

Las propiedades (*a*) y (*b*) son claras. La relación

(29) $$\psi_1 + \cdots + \psi_i = 1 - (1 - \varphi_1) \cdots (1 - \varphi_i)$$

es trivial para $i = 1$. Si se cumple (29) para algún $i < s$, la adición de (28) y (29) produce (29) con $i + 1$ en lugar de i. Se deduce que

(30) $$\sum_{i=1}^{s} \psi_i(\mathbf{x}) = 1 - \prod_{i=1}^{s} [1 - \varphi_i(\mathbf{x})] \qquad (\mathbf{x} \in R^n).$$

Si $\mathbf{x} \in K$, entonces $\mathbf{x} \in B(\mathbf{x}_i)$ para algún i, en consecuencia $\varphi_i(\mathbf{x}) = 1$, y el producto en (30) es 0. Esto demuestra (*c*).

CAMBIO DE VARIABLE

Se describirá ahora el efecto de un cambio de variables en una integral múltiple. Para facilitar esto, nos restringiremos a funciones continuas con soporte compacto, aunque esto es demasiado restrictivo para la mayoría de las aplicaciones. Esto se ilustra por medio de los Ejercicios 9 al 13.

10.9 Teorema *Supongamos que T es un mapeo-* $\mathscr{C}'$ 1-1 *de un conjunto abierto $E \subset R^k$ en R^k, tal que $J_T(\mathbf{x}) \neq 0$ para todo $\mathbf{x} \in E$. Si f es una función continua en R^k, cuyo soporte es compacto y está en $T(E)$, entonces*

(31) $$\int_{R^k} f(\mathbf{y})\, d\mathbf{y} = \int_{R^k} f(T(\mathbf{x}))|J_T(\mathbf{x})|\, d\mathbf{x}.$$

Recordemos que J_T es el Jacobiano de T. La hipótesis $J_T(\mathbf{x}) \neq 0$ implica por el teorema de la función inversa, que T^{-1} es continua en $T(E)$, lo que asegura que el integrando del segundo miembro de (31) tiene soporte compacto en E. (Teorema 4.14).

La aparición del *valor absoluto* de $J_T(\mathbf{x})$ en (31) requiere un comentario. Consideremos el caso $k = 1$, y supongamos que T es un mapeo- $\mathscr{C}'$ 1-1 de R^1 sobre R^1. Entonces $J_T(x) = T'(x)$, y si T es *creciente*, tenemos

$$(32) \qquad \int_{R^1} f(y)\,dy = \int_{R^1} f(T(x))T'(x)\,dx,$$

por los Teoremas 6.19 y 6.17, para toda f continua con soporte compacto. Pero si T decrece, será $T'(x) < 0$, y si f es positiva en el interior de su soporte, el primer miembro de (32) es positivo y el segundo negativo. Si se sustituye T' por $|T'|$ en (32), se obtiene una ecuación correcta.

Hasta ahora hemos considerado integrales de funciones sobre conjuntos de R^k sin asociar orientación ni dirección con estos subconjuntos. Cuando lleguemos a la integración de formas diferenciales sobre superficies, adoptaremos un punto de vista diferente.

Demostración De las observaciones que acabamos de hacer, se deduce que (31) es cierta si T es un mapeo- $\mathscr{C}'$ primitivo (ver la Definición 10.5) y el Teorema 10.2 demuestra también que (31) es cierta si T es una aplicación lineal que intercambia solamente dos coordenadas.

Si el teorema es cierto para las transformaciones P, Q y si $S(\mathbf{x}) = P(Q(\mathbf{x}))$, será

$$\begin{aligned}\int f(\mathbf{z})\,d\mathbf{z} &= \int f(P(\mathbf{y}))|J_P(\mathbf{y})|\,d\mathbf{y}\\ &= \int f(P(Q(\mathbf{x})))|J_P(Q(\mathbf{x}))|\,|J_Q(\mathbf{x})|\,d\mathbf{x}\\ &= \int f(S(\mathbf{x}))|J_S(\mathbf{x})|\,d\mathbf{x},\end{aligned}$$

puesto que

$$\begin{aligned}J_P(Q(\mathbf{x}))J_Q(\mathbf{x}) &= \det P'(Q(\mathbf{x}))\det Q'(\mathbf{x})\\ &= \det P'(Q(\mathbf{x}))Q'(\mathbf{x}) = \det S'(\mathbf{x}) = J_S(\mathbf{x}),\end{aligned}$$

por el teorema de la multiplicación de determinantes y la regla de la cadena. Así, pues, el teorema es cierto también para S.

Cada punto $\mathbf{a} \in E$ tiene una vecindad U, en la cual

$$(33) \qquad T(\mathbf{x}) = T(\mathbf{a}) + B_1 \cdots B_{k-1}\mathbf{G}_k \circ \mathbf{G}_{k-1} \circ \cdots \circ \mathbf{G}_1(\mathbf{x}-\mathbf{a}),$$

donde $\mathbf{G}_i$, B_i son como en el Teorema 10.7. Suponiendo que $V = T(U)$, se deduce que se cumple (31) si el soporte de $\mathbf{f}$ está en V. Así:

Todo punto $T(E)$ tiene una vecindad $V_\mathbf{y}$ tal que se cumple (31) *para toda f continua, cuyo soporte está en $V_\mathbf{y}$.*

Sea, ahora, f una función continua con soporte compacto $K \subset T(E)$. Como $\{V_\mathbf{y}\}$ cubre a K, el Corolario del Teorema 10.8 muestra que $f = \Sigma\psi_i f$, donde cada ψ_i es continua, y cada ψ_i tiene su soporte en

algún V_y. Por esto se cumple (31) para cada $\psi_i f$, y en consecuencia, también para su suma.

FORMAS DIFERENCIALES

Ahora se desarrollarán los mecanismos necesarios para la versión en n dimensiones del teorema fundamental del cálculo, que comúnmente se llama *teorema de Stokes*. La forma original del teorema de Stokes surge en aplicaciones del análisis vectorial al electromagnetismo y fue establecido en términos del rotacional de un campo vectorial. Otros casos especiales son el teorema de Green y el de la divergencia. Estos temas se estudian brevemente al final del capítulo.

Un detalle curioso del teorema de Stokes es que la única cosa difícil acerca de él, es la estructura elaborada de definiciones que son necesarias para su postulación. Estas definiciones se refieren a formas diferenciales, sus derivadas, fronteras y orientación. Una vez que se han entendido estos conceptos, el establecimiento del teorema es muy breve y conciso, y su demostración presenta muy poca dificultad.

Hasta, ahora, hemos considerado derivadas de funciones de varias variables solamente para funciones definidas en conjuntos *abiertos*, lo que se hizo por conveniencia para que nos permitiera soslayar las dificultades que se presentan en los puntos frontera. Sin embargo, es conveniente ahora estudiar las funcioes diferenciables en conjuntos *compactos*. En consecuencia, adoptaremos el siguiente convenio: Decir que $\mathbf{f}$ es un mapeo- $\mathscr{C}'$ (o un mapeo- $\mathscr{C}''$) de un conjunto compacto $D \subset R^k$ en R^n, significa que hay un mapeo- $\mathscr{C}'$ (o un mapeo- $\mathscr{C}''$) $\mathbf{g}$ de un conjunto abierto $W \subset R^k$ en R^n, tal que $D \subset W$ y que $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ para todo $\mathbf{x} \in D$.

10.10 Definición Supongamos que E es un conjunto abierto en R^n. Una *k-superficie en E* es un mapeo- $\mathscr{C}'$ Φ de un conjunto compacto $D \subset R^k$ en E.

A D se le llama *dominio de parámetros* de Φ. Representaremos los puntos de D por $\mathbf{u} = (u_1, \ldots, u_k)$.

Nos limitaremos al caso sencillo en que D es, o bien, una k-celda o el k-simplex Q^k descrito en el Ejemplo 10.4. La razón de esto es que tendremos que integrar sobre D, y todavía no hemos desarrollado una teoría de integración sobre conjuntos más complicados en R^k. Se verá que esta limitación a D (que haremos tácitamente desde ahora) no entraña pérdida de generalidad considerable en la teoría resultante de las formas diferenciales.

Nótese que las superficies-k en E están definidas como *mapeos* en E, no como subconjuntos de E. La comparación con la Definición 6.26 muestra que las superficies-1 no son más que curvas continuamente diferenciables.

10.11 Definición Supongamos que E es un conjunto abierto en R^n. *Una forma diferencial de orden* $k \geq 1$ *en E* (brevemente, una *k-forma en E*) es una función ω, representada simbólicamente por la suma

(34) $$\omega = \sum a_{i_1 \cdots i_k}(\mathbf{x})\, dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

(los índices $i_1, \ldots, i_k$ varían, independientemente, de 1 a n), que asigna a cada k-superficie Φ en E un número $\omega(\Phi) = \int_\Phi \omega$, de acuerdo con la regla

(35) $$\int_\Phi \omega = \int_D \sum a_{i_1 \cdots i_k}(\Phi(\mathbf{u})) \frac{\partial(x_{i_1}, \ldots, x_{i_k})}{\partial(u_1, \ldots, u_k)} d\mathbf{u},$$

donde D es el dominio de parámetros de Φ.

Las funciones $a_{i_1 \cdots i_k}$ se suponen reales y continuas en E. Si $\phi_1, \ldots, \phi_n$ son las componentes de Φ, el Jacobiano de (35) es el determinado por la aplicación

$$(u_1, \ldots, u_k) \rightarrow (\phi_{i_1}(\mathbf{u}), \ldots, \phi_{i_k}(\mathbf{u})).$$

Obsérvese que el segundo miembro de (35) es una integral sobre D, definida del modo hecho en la Definición 10.1 (o el Ejemplo 10.4) y que (35) es la *definición* del símbolo $\int_\Phi \omega$.

Se dice que una k-forma ω es la de clase $\mathscr{C}'$ o $\mathscr{C}''$ si las funciones $a_{i_1 \cdots i_k}$ en (34) son todas de clase $\mathscr{C}'$ o $\mathscr{C}''$.

Una 0-forma en E está definida como una función continua en E.

10.12 Ejemplos

(*a*) Sea γ una superficie-1 (una curva de clase $\mathscr{C}'$) en R^3, con dominio de parámetros [0, 1].

Si se escribe (x,y,z) en lugar de (x_1,x_2,x_3), y se pone

$$\omega = x\,dy + y\,dx.$$

Entonces

$$\int_\gamma \omega = \int_0^1 [\gamma_1(t)\gamma_2'(t) + \gamma_2(t)\gamma_1'(t)]\, dt = \gamma_1(1)\gamma_2(1) - \gamma_1(0)\gamma_2(0).$$

Nótese que en este ejemplo $\int_\gamma \omega$ depende solo del punto inicial $\gamma(0)$ y del punto final $\gamma(1)$ de γ. En particular, para cada curva cerrada γ se tiene $\int_\gamma \omega = 0$. (Como se verá después, esto es válido para cada 1-forma ω que sea *exacta*.)

A las integrales de las 1-formas se les llama comúnmente *integrales de línea*.

(*b*) Si con $a > 0$, $b > 0$ fijos, se define

$$\gamma(t) = (a \cos t, b \operatorname{sen} t) \qquad (0 \leq t \leq 2\pi),$$

de manera que γ es una curva cerrada en R^2. (Su rango es una elipse.) Entonces,

$$\int_\gamma x\,dy = \int_0^{2\pi} ab\cos^2 t\,dt = \pi ab,$$

mientras que

$$\int_\gamma y\,dx = -\int_0^{2\pi} ab\,\mathrm{sen}^2\,t\,dt = -\pi ab.$$

Nótese que $\int_\gamma x\,dy$ es el área de la región acotada por γ. Este es un caso especial del teorema de Green.

(c) Sea D la 3-celda definida por

$$0 \le r \le 1, \qquad 0 \le \theta \le \pi, \qquad 0 \le \varphi \le 2\pi.$$

Si se define $\Phi(r,\theta,\varphi) = (x,y,z)$, en donde

$$\begin{aligned} x &= r\,\mathrm{sen}\,\theta\cos\varphi \\ y &= r\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\varphi \\ z &= r\cos\theta. \end{aligned}$$

Entonces

$$J_\Phi(r,\theta,\varphi) = \frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)} = r^2\,\mathrm{sen}\,\theta.$$

De aquí que

(36) $$\int_\Phi dx \wedge dy \wedge dz = \int_D J_\Phi = \frac{4\pi}{3}.$$

Nótese que Φ mapea D sobre la bola unitaria cerrada de R^3, y que el mapeo es 1-1 en el interior de D (pero que algunos puntos frontera se identifican por medio de Φ), y finalmente, que la integral (36) es igual al volumen de $\Phi(D)$.

10.13 Propiedades elementales Sean ω, ω_1, ω_2 k-formas en E. Se escribe $\omega_1 = \omega_2$ si y solo si $\omega_1(\Phi) = \omega_2(\Phi)$ para cada k-superficie Φ en E. En particular, $\omega = 0$ significa que $\omega(\Phi) = 0$ para cada k-superficie Φ en E. Si c es un número real, entonces $c\omega$ es la k-forma definida por

(37) $$\int_\Phi c\omega = c\int_\Phi \omega,$$

y $\omega = \omega_1 + \omega_2$ significa que

(38) $$\int_\Phi \omega = \int_\Phi \omega_1 + \int_\Phi \omega_2$$

para cada k-superficie Φ en E. Nótese como caso especial de (37) que $-\omega$ se define de manera que

(39) $$\int_\Phi (-\omega) = -\int_\Phi d\omega.$$

Considérese una k-forma

(40) $$\omega = a(\mathbf{x})\, dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

y sea $\bar{\omega}$ la k-forma que se obtiene al intercambiar algún par de subíndices en (40). Si se combinan (35) y (39) teniendo en cuenta que un determinante cambia de signo si dos de sus renglones se intercambian, entonces se ve que

(41) $$\bar{\omega} = -\omega.$$

Como un caso especial de esto, nótese que se cumple la *relación anticonmutativa*

(42) $$dx_i \wedge dx_j = -dx_j \wedge dx_i$$

para todo i y j. En particular,

(43) $$dx_i \wedge dx_i = 0 \qquad (i = 1, \ldots, n).$$

De manera más general, volviendo a (40), y suponiendo que $i_r = i_s$ para algún $r \neq s$. Si estos dos subíndices se intercambian, entonces $\bar{\omega} = \omega$, de aquí que por (41), $\omega = 0$.

Dicho de otra manera, *si* ω *está dado por* (40), *entonces* $\omega = 0$ *a menos que los subíndices* $i_1, \ldots, i_n$ *sean todos distintos.*

Si ω está como en (34), pueden omitirse entonces los sumandos con subíndices repetidos sin cambiar ω.

Se deduce que 0 es la única k-forma en cualquier subconjunto abierto de R^n, si $k > n$.

La anticonmutatividad dada en (42) es la razón por la cual se prestó demasiada atención a los signos menos cuando se estudiaron formas diferenciables.

10.14 k-formas básicas Si $i_1, \ldots, i_k$ son enteros tales que $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, y si I es la k-ada ordenada $\{i_1, \ldots, i_k\}$, entonces I se llama un *k-índice creciente* y se usa la notación abreviada

(44) $$dx_I = dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

Estas formas dx_I son las llamadas *k-formas básicas* en R^n.

No es difícil verificar que hay exactamente $n!/k!(n-k)!$ k-formas básicas en R^n; no obstante, no se usarán.

Es mucho más importante el hecho de que cada k-forma puede representarse en términos de k-formas básicas. Para poder ver esto, nótese que cada k-tupla $\{j_1, \ldots, j_k\}$ de enteros distintos puede convertirse a un k-índice creciente J por medio de un número finito de intercambios de parejas; cada una de estas cantidades para una multiplicación por -1, como se vio en la sección 10.13; por consiguiente

(45) $$dx_{j_1} \wedge \cdots \wedge dx_{j_k} = \varepsilon(j_1, \ldots, j_k)\, dx_J$$

en donde $\epsilon(j_1, \ldots, j_k)$ es 1 o -1, dependiendo del número de intercambios que se necesiten. En efecto, se puede ver fácilmente que

(46) $$\varepsilon(j_1, \ldots, j_k) = s(j_1, \ldots, j_k)$$

en donde s es la de la Definición 9.33.

Por ejemplo,

$$dx_1 \wedge dx_5 \wedge dx_3 \wedge dx_2 = -dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_5$$

y

$$dx_4 \wedge dx_2 \wedge dx_3 = dx_2 \wedge dx_3 \wedge dx_4.$$

Si cada k-ada en (34) se convierte en un k-índice creciente, entonces se obtiene la llamada *representación estándar* de ω:

(47) $$\omega = \sum_I b_I(\mathbf{x})\, dx_I.$$

La sumatoria en (47) se extiende sobre todos los k-índices crecientes I. [Por supuesto, cada k-índice creciente surge de muchas (para ser precisos de $k!$) k-ada. Cada b_I en (47) puede ser entonces una suma de varios de los coeficientes que aparecen en (34).]

Por ejemplo,

$$x_1\, dx_2 \wedge dx_1 - x_2\, dx_3 \wedge dx_2 + x_3\, dx_2 \wedge dx_3 + dx_1 \wedge dx_2$$

es una 2-forma en R^3 cuya representación estándar es

$$(1 - x_1)\, dx_1 \wedge dx_2 + (x_2 + x_3)\, dx_2 \wedge dx_3.$$

Una de las principales razones para la introducción de la representación estándar de una k-forma la da el siguiente teorema de unicidad.

10.15 Teorema *Supóngase que*

(48) $$\omega = \sum_I b_I(\mathbf{x})\, dx_I$$

es la representación estándar de una k-forma ω *en un conjunto abierto* $E \subset R^n$. *Si* $\omega = 0$ *en E, entonces* $b_I(\mathbf{x}) = 0$ *para cada k-índice creciente I y para cada* $\mathbf{x} \in E$.

Nótese que la proposición análoga sería falsa para sumas como la (34), ya que, por ejemplo,

$$dx_1 \wedge dx_2 + dx_2 \wedge dx_1 = 0.$$

Demostración Para llegar a una contradicción, admitamos que $b_J(\mathbf{v}) > 0$ para algún $\mathbf{v} \in E$ y algún k-índice creciente $J = \{j_1, \ldots, j_k\}$. Como b_J es continua, existe $h > 0$ tal que $b_J(\mathbf{x}) > 0$ para todo $\mathbf{x} \in R^n$ cuyas coordenadas satisfagan $|x_i - v_i| \leq h$. Sea D la k-celda en R^k tal que $\mathbf{u} \in D$ si, y solo si $|u_r| \leq h$ para $r = 1, \ldots, k$. Si se define

(49) $$\Phi(\mathbf{u}) = \mathbf{v} + \sum_{r=1}^{k} u_r \mathbf{e}_{j_r} \qquad (\mathbf{u} \in D).$$

Entonces Φ es una k-superficie en E, con dominio de parámetros D, y $b_J(\Phi(\mathbf{u})) > 0$ para cada $\mathbf{u} \in D$.

Se requiere que

(50) $$\int_\Phi \omega = \int_D b_J(\Phi(\mathbf{u}))\, d\mathbf{u}.$$

Ya que el miembro derecho de (50) es positivo, se deduce que $\omega(\Phi) \neq$ $\neq 0$. Por consiguiente (50) proporciona la contradicción.

Aplíquese (35) a (48) para demostrar (50). Para ser más específicos, calcúlense los Jacobianos que aparecen en (35). De (49),

$$\frac{\partial(x_{j_1}, \ldots, x_{j_k})}{\partial(u_1, \ldots, u_k)} = 1.$$

Para cualquier otro k-índice creciente $I \neq J$, el Jacobiano es 0, porque es el determinante de una matriz que tiene al menos un renglón de ceros.

10.16 Productos de k-formas básicas Supóngase que

(51) $$I = \{i_1, \ldots, i_p\}, \qquad J = \{j_1, \ldots, j_q\}$$

donde $1 \leq i_1 < \cdots < i_p \leq n$ y $1 \leq j_1 < \cdots < j_q \leq n$. El *producto* de las formas básicas correspondientes dx_I y dx_J en R^n es una $(p + q)$-forma en R^n, representada por el símbolo $dx_I \wedge dx_J$, y definida por

(52) $$dx_I \wedge dx_J = dx_{i_1} \wedge \cdots \wedge dx_{i_p} \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_q}.$$

Si I y J tienen un elemento en común, entonces de la sección 10.13 se puede ver que $dx_I \wedge dx_J = 0$.

Si I y J no tienen elementos en común, se escribe $[I,J]$ para el $(p + q)$ -índice creciente que se obtiene ordenando los miembros de $I \cup J$ en forma creciente. Entonces $dx_{[I,J]}$ es una $(p + q)$-forma básica. Se requiere que

(53) $$dx_I \wedge dx_J = (-1)^\alpha dx_{[I,J]}$$

donde α es el número de diferencias $j_t - i_s$ que son *negativas.* (El número de diferencias positivas es entonces $pq - \alpha$.)

Para demostrar (53), se efectúan las siguientes operaciones en los números

(54) $$i_1, \ldots, i_p; j_1, \ldots, j_q.$$

Ahora se mueve i_p a la derecha, paso a paso, hasta que su vecino izquierdo sea menor que i_p. El número de pasos es el de subíndices t tales que $i_p < j_t$. (Nótese que una posibilidad distinta son 0 pasos.) En seguida se hace lo mismo para $i_{p-1}, \ldots, i_1$. El número total de pasos que se toman es α. El ordenamiento final que se alcanza es $[I,J]$. Cuando se aplica cada paso al miembro derecho de (52), multiplica $dx_I \wedge dx_J$ por -1. En consecuencia, se cumple (53).

Nótese que el miembro derecho de (53) es la representación estándar de $dx_I \wedge dx_J$.

En seguida, sea $K = (k_1, \ldots, k_r)$ un r-índice creciente en $\{1, \ldots, n\}$. Se usará (53) para demostrar que

(55) $$(dx_I \wedge dx_J) \wedge dx_K = dx_I \wedge (dx_J \wedge dx_K).$$

Si de los conjuntos I, J, K dos de ellos tienen un elemento en común, entonces cada miembro de (55) es 0; por lo tanto, son iguales.

Así que supóngase que I, J, K son ajenos por pares. Sea $[I, J, K]$ el $(p + q + r)$-índice creciente obtenido de sus uniones. Asóciese β con la pareja ordenada (J, K) y γ con la pareja ordenada (I, K) en la misma forma que α se asoció con $[I, J]$ en (53). Entonces el miembro izquierdo de (55) es, aplicando (53) dos veces,

$$(-1)^\alpha dx_{[I,J]} \wedge dx_K = (-1)^\alpha(-1)^{\beta+\gamma} dx_{[I,J,K]}$$

y el miembro derecho de (55) es

$$(-1)^{\beta}\, dx_I \wedge dx_{[J,K]} = (-1)^{\beta}(-1)^{\alpha+\gamma}\, dx_{[I,J,K]}.$$

En consecuencia (55) es correcta.

10.17 Multiplicación Supóngase que ω y λ son p- y q-formas, respectivamente, en algún conjunto abierto $E \subset R^n$, con representaciones estándar

$$\omega = \sum_I b_I(\mathbf{x})\, dx_I, \qquad \lambda = \sum_J c_J(\mathbf{x})\, dx_J \tag{56}$$

en donde I y J varían sobre todos los p-índices crecientes y sobre todos los q-índices crecientes tomados del conjunto $\{1, \ldots, n\}$.

Su producto, representado por el símbolo $\omega \wedge \lambda$, está definido como

$$\omega \wedge \lambda = \sum_{I,J} b_I(\mathbf{x})c_J(\mathbf{x})\, dx_I \wedge dx_J. \tag{57}$$

En esta suma, I y J varían independientemente sobre sus posibles valores, y $dx_I \wedge dx_J$ está dada como en la sección 10.16. Por esto $\omega \wedge \lambda$ es una $(p + q)$-forma en E.

Es fácil ver (los detalles se dejan como ejercicio) que las leyes distributivas

$$(\omega_1 + \omega_2) \wedge \lambda = (\omega_1 \wedge \lambda) + (\omega_2 \wedge \lambda)$$

y

$$\omega \wedge (\lambda_1 + \lambda_2) = (\omega \wedge \lambda_1) + (\omega \wedge \lambda_2)$$

se cumplen, con respecto a la adición definida en la sección 19.13. Si se combinan estas leyes distributivas con (55), se obtiene la ley asociativa

$$(\omega \wedge \lambda) \wedge \sigma = \omega \wedge (\lambda \wedge \sigma) \tag{58}$$

para formas arbitrarias ω, λ, σ en E.

En esta presentación se admitió tácitamente que $p \geq 1$ y $q \geq 1$. El producto de una 0-forma f con la p-forma ω dada por (56) se define sencillamente como la p-forma

$$f\omega = \omega f = \sum_I f(\mathbf{x}) b_I(\mathbf{x})\, dx_I.$$

Cuando f es una 0-forma, se acostumbra escribir $f\omega$, en lugar de $f \wedge \omega$.

10.18 Diferenciación Se definirá ahora un operador diferenciación d que asocia una $(k + 1)$-forma $d\omega$ con cada k-forma ω de clase $\mathscr{C}'$ en algún conjunto abierto $E \subset R^n$.

Una 0-forma de clase $\mathscr{C}'$ en E es una función real $f \in \mathscr{C}'(E)$, y se define

(59) $$df = \sum_{i=1}^{n} (D_i f)(\mathbf{x})\, dx_i .$$

Si $\omega = \Sigma b_I(\mathbf{x})\, dx_I$ es la presentación estándar de una k-forma ω, y $b_I \in \mathscr{C}'(E)$ para cada k-índice creciente I, entonces se define

(60) $$d\omega = \sum_{I} (db_I) \wedge dx_I .$$

10.19 Ejemplo Supóngase que E es abierto en $R^n f \in \mathscr{C}'(E)$, y γ es una curva continuamente diferenciable en E, con dominio $[0, 1]$. De (59) y (35) se tiene que

(61) $$\int_{\gamma} df = \int_0^1 \sum_{i=1}^{n} (D_i f)(\gamma(t))\gamma_i'(t)\, dt.$$

Por la regla de la cadena, el último integrando es $(f \circ \gamma)'(t)$. De aquí que

(62) $$\int_{\gamma} df = f(\gamma(1)) - f(\gamma(0)),$$

y se ve que $\int_\gamma dy$ es la misma para toda γ con el mismo punto inicial y final, de la misma manera que en (a) del Ejemplo 10.12.

Por tanto, al comparar con el Ejemplo 10.12(b), puede verse que la 1-forma $x\, dy$ no es la derivada de cualquier 0-forma f. Esto puede inferirse también de la parte (b) del siguiente teorema, debido a que

$$d(x\, dy) = dx \wedge dy \neq 0.$$

10.20 Teorema

(a) *Si ω y λ son k- y m-formas, respectivamente, de clase $\mathscr{C}'$ en E, entonces*

(63) $$d(\omega \wedge \lambda) = (d\omega) \wedge \lambda + (-1)^k\, \omega \wedge d\lambda.$$

(b) *Si ω es de clase $\mathscr{C}''$ en E, entonces $d^2\omega = 0$.*

Aquí, por supuesto que $d^2\omega$ significa $d(d\omega)$.

Demostración Debido a (57) y (60), se deduce (*a*) si se demuestra (63) para el caso especial

(64) $$\omega = f\,dx_I, \qquad \lambda = g\,dx_J$$

en donde $f, g \in \mathscr{C}'(E)$, dx_I es una k-forma básica, y dx_J es una m-forma básica. (Si k o m o ambas son 0, simplemente se omite dx_I o dx_J en (64); la siguiente demostración no se ve afectada por esto.) Entonces,

$$\omega \wedge \lambda = fg\,dx_I \wedge dx_J.$$

Admitamos que I y J no tienen elementos en común. [En otro caso, cada uno de los tres términos de (63) es 0.] Usando (53), se obtiene

$$d(\omega \wedge \lambda) = d(fg\,dx_I \wedge dx_J) = (-1)^{\alpha}\, d(fg\,dx_{[I,J]}).$$

De (59), $d(fg) = f\,dg + g\,df$. Por esto (60) nos da

$$\begin{aligned} d(\omega \wedge \lambda) &= (-1)^{\alpha}(f\,dg + g\,df) \wedge dx_{[I,J]} \\ &= (g\,df + f\,dg) \wedge dx_I \wedge dx_J. \end{aligned}$$

Como dg es una 1-forma y dx_I es una k-forma se obtiene

$$dg \wedge dx_I = (-1)^k\, dx_I \wedge dg,$$

de (42). En consecuencia

$$\begin{aligned} d(\omega \wedge \lambda) &= (df \wedge dx_I) \wedge (g\,dx_J) + (-1)^k (f\,dx_I) \wedge (dg \wedge dx_J) \\ &= (d\omega) \wedge \lambda + (-1)^k \omega \wedge d\lambda, \end{aligned}$$

lo cual demuestra (*a*).

Nótese que la ley asociativa (58) se usó con libertad.

Para demostrar la parte (*b*) considérese primero una 0-forma $f \in \mathscr{C}''$:

$$\begin{aligned} d^2 f &= d\left(\sum_{j=1}^{n} (D_j f)(\mathbf{x})\, dx_j\right) \\ &= \sum_{j=1}^{n} d(D_j f) \wedge dx_j \\ &= \sum_{i,j=1}^{n} (D_{ij} f)(\mathbf{x})\, dx_i \wedge dx_j. \end{aligned}$$

Como $D_{ij}f = D_{ji}f$ (por el Teorema 9.41) y $dx_i \wedge dx_j = -dx_j \wedge dx_i$, puede verse que $d^2f = 0$.

Si como en (64) $\omega = f\,dx_I$, entonces $d\omega = (df) \wedge dx_I$. Y por (60), $d(dx_I) = 0$.

Por lo anterior (63) muestra que

$$d^2\omega = (d^2f) \wedge dx_I = 0.$$

10.21 Cambio de variables Supóngase que E es un conjunto abierto en R^n, T es un mapeo- $\mathscr{C}'$ de E en un conjunto abierto $V \subset R^m$, y ω es una k-forma en V, cuya presentación estándar es

$$\omega = \sum_I b_I(\mathbf{y})\, dy_I. \tag{65}$$

(Se usará $\mathbf{y}$ para puntos de V, y $\mathbf{x}$ para puntos de E.)

Sean $t_1, \ldots, t_m$ las componentes de T: Si

$$\mathbf{y} = (y_1, \ldots, y_m) = T(\mathbf{x})$$

entonces $y_i = t_i(\mathbf{x})$. Como en (59),

$$dt_i = \sum_{j=1}^{n} (D_j t_i)(\mathbf{x})\, dx_j \qquad (1 \le i \le m). \tag{66}$$

Entonces cada dt_i es una 1-forma en E.

El mapeo T transforma ω en una k-forma ω_T en E, cuya definición es

$$\omega_T = \sum_I b_I(T(\mathbf{x}))\, dt_{i_1} \wedge \cdots \wedge dt_{i_k}. \tag{67}$$

En cada sumando de (67), $I = \{i_1, \ldots, i_k\}$ es un k-índice creciente.

El siguiente teorema muestra que la adición, multiplicación y diferenciación de formas están definidas de tal manera que conmutan con cambios de variables.

10.22 Teorema *Con E y T dadas como en la sección* 10.21, *sean ω y λ las k-y m-formas en V, respectivamente. Entonces*

(*a*) $(\omega + \lambda)_T = \omega_T + \lambda_T$ *si* $k = m$;
(*b*) $(\omega \wedge \lambda)_T = \omega_T \wedge \lambda_T$;
(*c*) $d(\omega_T) = (d\omega)_T$ *si ω es de clase $\mathscr{C}'$ y T es de clase $\mathscr{C}''$.*

Demostración De las definiciones se deduce inmediatamente la parte (*a*). La parte (*b*) es casi obvia, una vez que se da uno cuenta que

$$(dy_{i_1} \wedge \cdots \wedge dy_{i_r})_T = dt_{i_1} \wedge \cdots \wedge dt_{i_r} \tag{68}$$

sin importar si $\{i_1, \ldots, i_r\}$ es creciente o no; (68) es válida porque se necesita el mismo número de signos menos en cada miembro de (68) para producir los reordenamientos crecientes.

Volviendo a la demostración de (c), si f es una 0-forma de clase $\mathscr{C}'$ en V, entonces

$$f_T(\mathbf{x}) = f(T(\mathbf{x})), \qquad df = \sum_i (D_i f)(\mathbf{y})\, dy_i.$$

de la regla de la cadena se deduce que

(69)
$$\begin{aligned} d(f_T) &= \sum_j (D_j f_T)(\mathbf{x})\, dx_j \\ &= \sum_j \sum_i (D_i f)(T(\mathbf{x}))(D_j t_i)(\mathbf{x})\, dx_j \\ &= \sum_i (D_i f)(T(\mathbf{x}))\, dt_i \\ &= (df)_T. \end{aligned}$$

Si $dy_I = dy_{i_1} \wedge \cdots \wedge dy_{i_k}$, entonces $(dy_I)_T = dt_{i_1} \wedge \cdots \wedge dt_{i_k}$, y el Teorema 10.20 muestra que

(70)
$$d((dy_I)_T) = 0.$$

(Aquí es donde se usa la suposición $T \in \mathscr{C}''$.)

Admítase ahora que $\omega = f\, dy_I$. Entonces

$$\omega_T = f_T(\mathbf{x})\, (dy_I)_T$$

y los cálculos anteriores conducen a

$$\begin{aligned} d(\omega_T) &= d(f_T) \wedge (dy_I)_T = (df)_T \wedge (dy_I)_T \\ &= ((df) \wedge dy_I)_T = (d\omega)_T. \end{aligned}$$

De (63) y (70) se ve que la primera igualdad se cumple, la segunda se cumple por (69), la tercera por la parte (b) y la última a partir de la definición de $d\omega$.

El caso general de (c) se deduce del caso especial ya demostrado, si se aplica (a). Esto completa la demostración.

Nuestro siguiente objetivo es el Teorema 10.25. Éste se deducirá directamente a partir de otras dos propiedades importantes de transformación de formas diferenciales, que se establecerán primero.

10.23 Teorema *Supongamos que T es mapeo- $\mathscr{C}'$ de un conjunto abierto $E \subset R^n$ en un conjunto abierto $V \subset R^m$, S es un mapeo- $\mathscr{C}'$ de V en un conjunto abierto $W \subset R^p$, y ω es una k-forma en W, de modo que ω_S es una k-forma en V y lo mismo $(\omega_S)_T$ que ω_{ST} son k-formas en E, donde ST está definida por $(ST)(\mathbf{x}) = S(T(\mathbf{x}))$. Entonces*

$$(\omega_S)_T = \omega_{ST}. \tag{71}$$

Demostración Si ω y λ son formas en W, el Teorema 10.22 muestra que

$$((\omega \wedge \lambda)_S)_T = (\omega_S \wedge \lambda_S)_T = (\omega_S)_T \wedge (\lambda_S)_T$$

y

$$(\omega \wedge \lambda)_{ST} = \omega_{ST} \wedge \lambda_{ST}.$$

Así pues, si se cumple (71) para ω y λ, se deduce que también se cumple para $\omega \wedge \lambda$. Como cada forma se puede construir a partir de 0-formas y 1-formas por adición y multiplicación, y como (71) es trivial para las 0-formas, es suficiente probar (71) en el caso $\omega = dz_q$, $q = 1, \ldots, p$. (Representaremos los puntos de E, V, W, por $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, respectivamente.)

Sean $t_1, \ldots, t_m$ los componentes de T, $s_1, \ldots, s_p$ los componentes de S, y $r_1, \ldots, r_p$ los componentes de ST. Si $\omega = dz_q$, será

$$\omega_S = ds_q = \sum_j (D_j s_q)(\mathbf{y})\, dy_j,$$

de modo que la regla de la cadena implica

$$\begin{aligned}(\omega_S)_T &= \sum_j (D_j s_q)(T(\mathbf{x}))\, dt_j \\ &= \sum_j (D_j s_q)(T(\mathbf{x})) \sum_i (D_i t_j)(\mathbf{x})\, dx_i \\ &= \sum_i (D_i r_q)(\mathbf{x})\, dx_i = dr_q = \omega_{ST}.\end{aligned}$$

10.24 Teorema *Supongamos que ω es una k-forma en un conjunto abierto $E \subset R^n$, Φ es una k-superficie en E, con dominio de parámetros $D \subset R^k$, y Δ es la k-superficie en R^k, con dominio de parámetros D definido por $\Delta(\mathbf{u}) = \mathbf{u}(\mathbf{u} \in D)$. Entonces*

$$\int_\Phi \omega = \int_\Delta \omega_\Phi.$$

Demostración Tenemos que considerar solamente el caso

$$\omega = a(\mathbf{x})\, dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

Si $\phi_1, \ldots, \phi_n$ son los componentes de Φ, entonces

$$\omega_\Phi = a(\Phi(\mathbf{u}))\, d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k}.$$

Se deducirá el teorema, si podemos demostrar que

(72) $$d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k} = J(\mathbf{u})\, du_1 \wedge \cdots \wedge du_k,$$

donde

$$J(\mathbf{u}) = \frac{\partial(x_{i_1}, \ldots, x_{i_k})}{\partial(u_1, \ldots, u_k)},$$

pues (72) implica

$$\int_\Phi \omega = \int_D a(\Phi(\mathbf{u}))J(\mathbf{u})\, d\mathbf{u}$$
$$= \int_\Delta a(\Phi(\mathbf{u}))J(\mathbf{u})\, du_1 \wedge \cdots \wedge du_k = \int_\Delta \omega_\Phi.$$

Sea $[A]$ la matriz k por k con elementos

$$\alpha(p, q) = (D_q \phi_{i_p})(\mathbf{u}) \qquad (p, q = 1, \ldots, k).$$

Entonces,

$$d\phi_{i_p} = \sum_q \alpha(p, q)\, du_q$$

de modo que

$$d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k} = \sum \alpha(1, q_1) \cdots \alpha(k, q_k)\, du_{q_1} \wedge \cdots \wedge du_{q_k}.$$

En esta última suma, $q_1, \ldots, q_k$, varían independientemente sobre $1, \ldots, k$. La relación anticonmutativa (42) implica que

$$du_{q_1} \wedge \cdots \wedge du_{q_k} = s(q_1, \ldots, q_k)\, du_1 \wedge \cdots \wedge du_k,$$

donde s es como en la Definición 9.33; aplicando esta definición vemos que

$$d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k} = \det\,[A]\, du_1 \wedge \cdots \wedge du_k;$$

y como $J(\mathbf{u}) = \det\,[A]$, queda demostrada (72).

La conclusión final de esta sección combina los dos teoremas precedentes.

10.25 Teorema *Supóngase que T es un mapeo- $\mathscr{C}'$ de un conjunto abierto $E \subset R^n$ en un conjunto abierto $V \subset R^m$, Φ es una k-superficie en E, y ω es una k-forma en V.*

Entonces,

$$\int_{T\Phi} \omega = \int_{\Phi} \omega_T .$$

Demostración Sea D el dominio de parámetros de Φ (por esto también de $T\Phi$) y Δ definida como en el Teorema 10.24.

Entonces,

$$\int_{T\Phi} \omega = \int_{\Delta} \omega_{T\Phi} = \int_{\Delta} (\omega_T)_\Phi = \int_{\Phi} \omega_T .$$

La primera de estas desigualdades es el Teorema 10.24, aplicado a $T\Phi$ en lugar de Φ. La segunda se deduce del Teorema 10.23. La tercera es el Teorema 10.24, con ω_T en lugar de ω.

CADENAS Y SIMPLEX

10.26 Simplex afines Un mapeo $\mathbf{f}$ que lleva un espacio vectorial X en un espacio vectorial Y se dice que es *afín* si $\mathbf{f} - \mathbf{f}(\mathbf{0})$ es lineal. En otras palabras, se requiere que

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{0}) + A\mathbf{x} \tag{73}$$

para algún $A \in L(X, Y)$.

Entonces un mapeo afín de R^k en R^n se determina si se conocen $\mathbf{f}(\mathbf{0})$ y $\mathbf{f}(\mathbf{e}_i)$ para $1 \leq i \leq k$; como de costumbre, $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ es la base estándar de R^k.

Se define el *simplex estándar* como el conjunto Q^k de todos los $\mathbf{u} \in R^k$ de la forma

$$\mathbf{u} = \sum_{i=1}^{k} \alpha_i \mathbf{e}_i \tag{74}$$

tales que $\alpha_i \geq 0$ para $i = 1, \ldots, k$ y $\Sigma\alpha_i \leq 1$.

Supóngase ahora que $\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_k$ son puntos de R^n. Se define el *k-simplex afín orientado*

$$\sigma = [\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_k] \tag{75}$$

como la k-superficie en R^n con dominio de parámetros Q^k dado por el mapeo afín

(76) $$\sigma(\alpha_1\mathbf{e}_1 + \cdots + \alpha_k\mathbf{e}_k) = \mathbf{p}_0 + \sum_{i=1}^{k}\alpha_i(\mathbf{p}_i - \mathbf{p}_0).$$

Nótese que σ se caracteriza por

(77) $$\sigma(\mathbf{0}) = \mathbf{p}_0, \qquad \sigma(\mathbf{e}_i) = \mathbf{p}_i \qquad (\text{para } 1 \leq i \leq k),$$

y que

(78) $$\sigma(\mathbf{u}) = \mathbf{p}_0 + A\mathbf{u} \qquad (\mathbf{u} \in Q^k)$$

en donde $A \in L(R^k, R^n)$ y $A\mathbf{e}_i = \mathbf{p}_i - \mathbf{p}_0$ para $1 \leq i \leq k$.

Llamamos a σ *orientado* para recalcar que se ha tenido en cuenta el orden de los vértices $\mathbf{p}_0, \ldots, \mathbf{p}_k$. Si

(79) $$\bar{\sigma} = [p_{i_0}, p_{i_1}, \ldots, p_{i_k}],$$

donde $\{i_0, i_1, \ldots, i_k\}$ es una permutación del conjunto ordenado $\{0, 1, \ldots, k\}$ adoptaremos la notación

(80) $$\bar{\sigma} = s(i_0, i_1, \ldots, i_k)\sigma,$$

siendo s la función definida en la Definición 9.33. Así que $\bar{\sigma} = \pm\,\sigma$, dependiendo de si $s = 1$ o $s = -1$. Hablando con rigor, habiendo adoptado (75), y (76) como definición de σ, no escribiríamos $\bar{\sigma} = \sigma$ si no fuera $i_0 = 0, \ldots, i_k = k$, aun cuando $s(i_0, \ldots, i_k) = 1$; lo cual consideramos ahora es una relación de equivalencia, no una igualdad. Sin embargo, para nuestro objeto, se justifica la notación, por el Teorema 10.27.

Si $\bar{\sigma} = \varepsilon\sigma$ (utilizando el convenio anterior) y si $\varepsilon = 1$, decimos que $\bar{\sigma}$ y σ tiene la *misma orientación*; si $\varepsilon = -1$, se dice que $\bar{\sigma}$ y σ tienen *orientaciones opuestas*. Obsérvese que no hemos definido lo que entendemos por «orientación de un simplex». Lo que hemos definido es una relación entre pares de simplex que tienen el mismo conjunto de vértices, siendo esta relación la de «tener la misma orientación».

Sin embargo, existe una situación en la cual la orientación de un simplex se puede definir de manera natural. Esta ocurre cuando $n = k$ y los vectores $\mathbf{p}_i - \mathbf{p}_0$ $(1 \leq i \leq k)$ son *independientes*. En ese caso, la transformación lineal A que aparece en (78) es invertible, y su determinante (que es el mismo que el Jacobiano de σ) no es 0. Se dice entonces que σ está orientado *positivamente* (o *negativamente*) si det A es positivo (o negativo). En particular, el simplex $[\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_k]$ en R^k, dado por el mapeo identidad, tiene orientación positiva.

Hasta ahora hemos supuesto $k \geq 1$. Un 0-*simplex orientado* se define como un punto al que se le atribuye un signo. Escribimos $\sigma = +\mathbf{p}_0$ o $\sigma =$

$= -\mathbf{p}_0$. Si $\sigma = \varepsilon\mathbf{p}_0(\varepsilon = \pm 1)$, y f es una 0-forma (esto es, una función real), definiremos

$$\int_\sigma f = \varepsilon f(p_0).$$

10.27 Teorema *Si σ es un k-simplex rectilíneo orientado en un conjunto abierto $E \subset R^n$ y si $\bar{\sigma} = \varepsilon\sigma$,*

(81)
$$\int_{\bar{\sigma}} \omega = \varepsilon \int_\sigma \omega$$

para toda k-forma ω en E.

Demostración Para $k = 0$, de la anterior definición, se deduce (81). Así, admitiremos que $k \geq 1$ y que σ está dada por (75).

Supongamos que $1 < j \leq k$, y que se obtiene $\bar{\sigma}$ de σ permutando $\mathbf{p}_0$ y $\mathbf{p}_j$. Entonces, $\varepsilon = -1$, y

$$\bar{\sigma}(\mathbf{u}) = \mathbf{p}_j + B\mathbf{u} \qquad (\mathbf{u} \in Q^k),$$

donde B es el mapeo lineal de R^k en R^n definido por $B\mathbf{e}_j = \mathbf{p}_0 - \mathbf{p}_j$, $B\mathbf{e}_i = \mathbf{p}_i - \mathbf{p}_j$, si $i \neq j$. Si escribimos $A\mathbf{e}_i = \mathbf{x}_i$ $(1 \leq i \leq k)$, donde A está dado por (78), los vectores columna de B (esto es, los vectores $B\mathbf{e}_i$) son

$$\mathbf{x}_1 - \mathbf{x}_j, \ldots, \mathbf{x}_{j-1} - \mathbf{x}_j, -\mathbf{x}_j, \mathbf{x}_{j+1} - \mathbf{x}_j, \ldots, \mathbf{x}_k - \mathbf{x}_j.$$

Si restamos la columna j-ésima de cada una de las otras, ninguno de los determinantes (35) queda afectado, y obtendremos columnas $\mathbf{x}_1, \ldots, \mathbf{x}_{j-1}, -\mathbf{x}_j, \mathbf{x}_{j+1}, \ldots, \mathbf{x}_k$, que defieren de las de A solamente en el signo de la columna j-ésima. Por tanto, en este caso se cumple (81).

Supongamos, ahora, que $0 < i < j \leq k$ y que $\bar{\sigma}$ se obtiene de σ permutando $\mathbf{p}_i$ y $\mathbf{p}_j$. Entonces, $\bar{\sigma}(\mathbf{u}) = \mathbf{p}_0 + C\mathbf{u}$, donde C tiene las mismas columnas que A, excepto que se han permutado la j-ésima y la i-ésima; esto implica de nuevo, que se cumple (81), pues $\varepsilon = -1$.

El caso general se deduce teniendo en cuenta que toda permutación de $\{0, 1, \ldots, k\}$ es una combinación de los casos particulares que hemos considerado.

10.28 Definición *Una k-cadena afín* Γ en un conjunto abierto $E \subset R^n$ es una colección de un número finito de k-simplices afines orientados $\sigma_1, \ldots, \sigma_r$ en E no necesariamente distintos; en Γ puede ocurrir un simplex con alguna multiplicidad.

Si Γ tiene el significado anterior, y si ω es una k-forma en E, definimos

(82) $$\int_{\Gamma} \omega = \sum_{i=1}^{r} \int_{\sigma_i} \omega.$$

Podemos considerar una k-superficie Φ en E como una función cuyo dominio de definición es la colección de todas las k-formas ω en E, que asigna el número $\int_{\Phi} \omega$ a ω. Las funciones de variables reales pueden sumarse (ver la Definición 4.3), por lo que (82) sugiere la notación

(83) $$\Gamma = \sigma_1 + \cdots + \sigma_r$$

o, en forma reducida,

(84) $$\Gamma = \sum_{i=1}^{r} \sigma_i$$

para establecer el hecho de que (82) se cumple para cada k-forma ω en E.

Para evitar confusiones, las notaciones que se introdujeron en (83) y (80) tienen que manejarse con cuidado. El hecho es que cada k-simplex afín orientado σ en R^n es una función de dos maneras con diferentes dominios y rangos, y que por tanto, son posibles dos operaciones de adición completamente diferentes. Se definió σ originalmente como una función valuada en R^n con dominio Q^k; debido a esto $\sigma_1 + \sigma_2$ podría interpretarse como la función σ que asigna el vector $\sigma_1(\mathbf{u}) + \sigma_2(\mathbf{u})$ a cada $\mathbf{u} \in Q^k$; nótese que entonces, otra vez σ es un k-simplex afín orientado en R^n! Esto *no* es lo que debe entenderse por (83).

Por ejemplo, si como en (80) $\sigma_2 = -\sigma_1$ (es decir, si σ_1 y σ_2 tienen el mismo conjunto de vértices pero están orientados opuestamente) y si $\Gamma = \sigma_1 + \sigma_2$, entonces $\int_{\Gamma} \omega = 0$ para todo ω, y puede expresarse esto escribiendo $\Gamma = 0$ o $\sigma_1 + \sigma_2 = 0$. Esto no quiere decir que $\sigma_1(\mathbf{u}) + \sigma_2(\mathbf{u})$ es el vector nulo de R^n.

10.29 Fronteras Para $k \geq 1$, la *frontera* del k-simplex afín orientado

$$\sigma = [\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_k]$$

se define como la $(k-1)$-cadena afín

(85) $$\partial\sigma = \sum_{j=0}^{k} (-1)^j [\mathbf{p}_0, \ldots, \mathbf{p}_{j-1}, \mathbf{p}_{j+1}, \ldots, \mathbf{p}_k].$$

Por ejemplo, si $\sigma = [\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2]$, entonces

$$\partial\sigma = [\mathbf{p}_1, \mathbf{p}_2] - [\mathbf{p}_0, \mathbf{p}_2] + [\mathbf{p}_0, \mathbf{p}_1] = [\mathbf{p}_0, \mathbf{p}_1] + [\mathbf{p}_1, \mathbf{p}_2] + [\mathbf{p}_2, \mathbf{p}_0],$$

que coincide con la noción usual de la frontera orientada de un triángulo.

Obsérvese que, para $1 \leq j \leq k$, el simplex $\sigma_j = [\mathbf{p}_0, \ldots, \mathbf{p}_{j-1}, \mathbf{p}_{j+1}, \ldots, p_k]$ que aparece en (85) tiene como dominio de parámetros a Q^{k-1} y está definido como

(86) $$\sigma_j(\mathbf{u}) = \mathbf{p}_0 + B\mathbf{u} \qquad (\mathbf{u} \in Q^{k-1}),$$

en donde B es el mapeo lineal de R^{k-1} a R^n determinado por

$$B\mathbf{e}_i = \mathbf{p}_i - \mathbf{p}_0 \qquad (\text{si} \quad 1 \leq i \leq j-1),$$
$$B\mathbf{e}_i = \mathbf{p}_{i+1} - \mathbf{p}_0 \qquad (\text{si} \quad j \leq i \leq k-1).$$

El simplex

$$\sigma_0 = [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k],$$

que también aparece en (85), está dado por el mapeo

$$\sigma_0(\mathbf{u}) = \mathbf{p}_1 + B\mathbf{u},$$

en donde $B\mathbf{e}_i = \mathbf{p}_{i+1} - \mathbf{p}_1$ para $1 \leq i \leq k-1$.

10.30 Cadenas y simplex diferenciables Sea T un mapeo- $\mathscr{C}''$ de un conjunto abierto $E \subset R^n$ en un conjunto abierto $V \subset R^m$; no necesita ser T uno-a-uno. Si σ es un k-simplex afín orientado en E, entonces el mapeo compuesto $\Phi = T \circ \sigma$ (que algunas veces se escribirá en la forma simple $T\sigma$) es una k-superficie en V, con dominio de parámetros Q^k. A Φ se le llama *k-simplex orientado de clase* $\mathscr{C}''$.

Una colección finita Ψ de k-simplex orientados $\Phi_1, \ldots, \Phi_r$ de clase $\mathscr{C}''$ en V se llama una *k-cadena de clase* $\mathscr{C}''$ en V. Si ω es una k-forma en V se define

(87) $$\int_\Psi \omega = \sum_{i=1}^{r} \int_{\Phi i} \omega$$

y se usa la notación $\Psi = \Sigma\Phi_i$.

Si $\Gamma = \Sigma\sigma_i$ es una cadena afín y $\Phi_i = T \circ \sigma_i$, se define también $\Psi = T \circ \Gamma$, o

(88) $$T(\sum \sigma_i) = \sum T\sigma_i.$$

La frontera $\partial\Phi$ del k-simplex orientado $\Phi = T \circ \sigma$, se define como la $(k-1)$-cadena

(89) $$\partial\Phi = T(\partial\sigma).$$

Para justificar (89), obsérvese que si T es afín, entonces $\Phi = T \circ \sigma$ es un k-simplex afín orientado en el cual el caso (89) no compete a la definición, pero se ve que es *consecuencia* de (85). De este modo (89) generaliza este caso especial.

Si lo de Φ es verdadero, es inmediato que $\partial\Phi$ es de clase $\mathscr{C}''$.

Se define finalmente, la frontera $\partial\Psi$ de la k-cadena $\Psi = \Sigma\Phi_i$ como la $(k - 1)$-cadena

(90) $$\partial\Psi = \sum \partial\Phi_i .$$

10.31 Fronteras orientadas positivamente Hasta ahora se han asociado fronteras con cadenas, no con subconjuntos de R^n. Esta noción de frontera es exactamente la más adecuada para establecer y demostrar el teorema de Stokes. No obstante, es más conveniente y se acostumbra también en aplicaciones, especialmente en R^2 o R^3, hablar de "fronteras orientadas" de determinados conjuntos. Ahora se describirá brevemente esto.

Sea Q^n el simplex estándar en R^n, y σ_0 el mapeo identidad con dominio Q^n. Como se vio en la sección 10.26, σ_0 puede considerarse como un n-simplex positivamente orientado en R^n. Su frontera $\partial\sigma_0$ es una $(n - 1)$-cadena afín. Esta cadena se denomina *frontera positivamente orientada del conjunto* Q^n.

Por ejemplo, la frontera orientada positivamente de Q^3 es

$$[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3] - [0, \mathbf{e}_2, \mathbf{e}_3] + [0, \mathbf{e}_1, \mathbf{e}_3] - [0, \mathbf{e}_1, \mathbf{e}_2].$$

Sea ahora T un mapeo 1-1 de Q^n en R^n, de clase $\mathscr{C}''$, cuyo Jacobiano es (al menos en el interior de Q^n) positivo. Sea $E = T(Q^n)$. Por el teorema de la función inversa, E es la cerradura de un subconjunto abierto de R^n. Se define la frontera orientada positivamente del conjunto E como la $(n - 1)$-cadena

$$\partial T = T(\partial\sigma_0),$$

y esta $(n - 1)$-cadena puede representarse con ∂E.

Surge aquí una pregunta obvia: Si $E = T_1(Q^n) = T_2(Q^n)$ y si T_1 y T_2 tienen Jacobianos positivos, ¿es cierto que $\partial T_1 = \partial T_2$?; es decir, ¿es válida la desigualdad

$$\int_{\partial T_1} \omega = \int_{\partial T_2} \omega$$

para cada $(n - 1)$-forma ω? La respuesta es sí, pero se omitirá la demostración. (Como ejemplo, compárese el final de esta sección, con el Ejercicio 17).

Ampliando lo anterior, sea

$$\Omega = E_1 \cup \cdots \cup E_r,$$

donde $E_i = T_i(Q^n)$, cada T_i tiene las propiedades que tuvo T, y los interiores de los conjuntos E_i son ajenos por parejas. Entonces la $(n - 1)$-cadena

$$\partial T_1 + \cdots + \partial T_r = \partial \Omega$$

se denomina frontera orientada positivamente de Ω.

Por ejemplo, el cuadrado unitario I^2 en R^2 es la unión de $\sigma_1(Q^2)$ y $\sigma_2(Q^2)$, donde

$$\sigma_1(\mathbf{u}) = \mathbf{u}, \qquad \sigma_2(\mathbf{u}) = \mathbf{e}_1 + \mathbf{e}_2 - \mathbf{u}.$$

Ambos, σ_1 y σ_2 tienen Jacobiano $1 > 0$. Debido a que

$$\sigma_1 = [\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2], \qquad \sigma_2 = [\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1]$$

se tiene

$$\begin{aligned} \partial\sigma_1 &= [\mathbf{e}_1, \mathbf{e}_2] - [\mathbf{0}, \mathbf{e}_2] + [\mathbf{0}, \mathbf{e}_1], \\ \partial\sigma_2 &= [\mathbf{e}_2, \mathbf{e}_1] - [\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1] + [\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_2]; \end{aligned}$$

La suma de estas dos fronteras es

$$\partial I^2 = [\mathbf{0}, \mathbf{e}_1] + [\mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2] + [\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_2] + [\mathbf{e}_2, \mathbf{0}],$$

es decir, la frontera orientada positivamente de I^2. Nótese que $[\mathbf{e}_1, \mathbf{e}_2]$ canceló a $[\mathbf{e}_2, \mathbf{e}_1]$.

Si Φ es una 2-superficie en R^m, con dominio de parámetros I^2, entonces Φ (considerada como una función sobre 2-formas) es la misma que la 2-cadena

$$\Phi \circ \sigma_1 + \Phi \circ \sigma_2.$$

Entonces

$$\begin{aligned} \partial\Phi &= \partial(\Phi \circ \sigma_1) + \partial(\Phi \circ \sigma_2) \\ &= \Phi(\partial\sigma_1) + \Phi(\partial\sigma_2) = \Phi(\partial I^2). \end{aligned}$$

En otras palabras, si el dominio de parámetros de Φ es el cuadrado I^2, no se necesita hacer referencia al simplex Q^2, pero puede obtenerse $\partial\Phi$ directamente de ∂I^2.

Se pueden encontrar otros ejemplos en los ejercicios 17 a 19.

10.32 Ejemplo Si se define para $0 \leq u \leq \pi$, $0 \leq v \leq 2\pi$

$$\Sigma(u, v) = (\operatorname{sen} u \cos v, \operatorname{sen} u \operatorname{sen} v, \cos u).$$

Entonces Σ es una 2-superficie en R^3, cuyo dominio de parámetros es un rectángulo $D \subset R^2$, y cuyo rango es la esfera unitaria en R^3. Su frontera es

$$\partial\Sigma = \Sigma(\partial D) = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$$

en donde

$$\begin{aligned}
\gamma_1(u) &= \Sigma(u, 0) = (\operatorname{sen} u, 0, \cos u),\\
\gamma_2(v) &= \Sigma(\pi, v) = (0, 0, -1),\\
\gamma_3(u) &= \Sigma(\pi - u, 2\pi) = (\operatorname{sen} u, 0, -\cos u),\\
\gamma_4(v) &= \Sigma(0, 2\pi - v) = (0, 0, 1),
\end{aligned}$$

con $[0,\pi]$ y $[0, 2\pi]$ como intervalos de parámetros para u y v, respectivamente.

Como γ_2 y γ_4 son constantes, sus derivadas son 0, por esto la integral de cualquier 1-forma sobre γ_2 o γ_4 es 0. [Véase Ejemplo 1.12(a).]

Como $\gamma_3(u) = \gamma_1(\pi - u)$, la aplicación directa de (35) muestra que

$$\int_{\gamma_3} \omega = -\int_{\gamma_1} \omega$$

para cada 1-forma ω. Entonces $\int_{\partial\Sigma} \omega = 0$, y se concluye que $\partial\Sigma = 0$.

(En terminología geográfica, $\partial\Sigma$ parte del polo norte N, corre hacia el polo sur S a lo largo de un meridiano, se detiene en S, regresa a N por el mismo meridiano y finalmente se detiene en N. Los dos viajes por el meridiano son en direcciones opuestas. Por tanto, las dos integrales de línea correspondientes se cancelan una con otra. En el Ejercicio 32, hay también una curva que aparece dos veces en la frontera, pero sin cancelación.)

TEOREMA DE STOKES

10.33 Teorema *Si Ψ es una k-cadena de clase $\mathscr{C}''$ en un conjunto abierto $V \subset R^m$ y si ω es una $(k - 1)$forma de clase $\mathscr{C}'$ en V, entonces*

$$\int_{\Psi} d\omega = \int_{\partial\Psi} \omega. \tag{91}$$

El caso $k = m = 1$ es nada menos que el teorema fundamental del cálculo (con una suposición adicional de diferenciabilidad). El caso $k = m = 2$ es el teorema de Green, y $k = m = 3$ da el llamado "teorema de la divergencia" de Gauss. El caso $k = 2$, $m = 3$ es el que originalmente des-

cubrió Stokes. (En el libro de Spivak se reseñan algunos antecedentes históricos.) Estos casos especiales se estudiarán al final de este capítulo.

Demostración Es suficiente con demostrar que

(92) $$\int_{\Phi} d\omega = \int_{\partial\Phi} \omega$$

para cada k-simplex Φ orientado de clase $\mathscr{C}''$ en V.

Pero si (92) se demuestra, y si $\Psi = \Sigma\Phi_i$, entonces (87) y (89) implican (91).

Fíjese tal Φ y hágase

(93) $$\sigma = [\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_k].$$

Por lo anterior, σ es el k-simplex afín orientado con dominio de parámetro Q^k definido por el mapeo identidad. Como Φ está también definido sobre Q^k (véase Definición 10.30) y $\Phi \in \mathscr{C}''$, hay un conjunto abierto $E \subset R^k$ que contiene a Q^k, y un mapeo- $\mathscr{C}''T$ de E en V tal que $\Phi = T \circ \sigma$. Por los Teoremas 10.25 y 10.22(c), el miembro izquierdo de (92) es igual a

$$\int_{T\sigma} d\omega = \int_{\sigma} (d\omega)_T = \int_{\sigma} d(\omega_T).$$

Otra aplicación del Teorema 10.25 muestra, por (89), que el miembro derecho de (92) es

$$\int_{\partial(T\sigma)} \omega = \int_{T(\partial\sigma)} \omega = \int_{\partial\sigma} \omega_T .$$

Como ω_T es una $(k - 1)$-forma en E, se ve que *para demostrar* (92) *es suficiente con mostrar que*

(94) $$\int_{\sigma} d\lambda = \int_{\partial\sigma} \lambda$$

para el simplex especial (93) *y para cada* $(k - 1)$*-forma* λ *de clase* $\mathscr{C}'$ *en* E.

Si $k = 1$, la definición de un 0-simplex orientado muestra que (94) tan sólo asegura que

(95) $$\int_0^1 f'(u)\, du = f(1) - f(0)$$

para cada función f continuamente diferenciable sobre [0, 1], lo cual es verdadero por el teorema fundamental del cálculo.

De ahora en adelante se supondrá que $k > 1$, con un entero r ($1 \leq r \leq k$) fijo, y se elige $f \in \mathscr{C}'(E)$. Entonces basta con probar (94) para el caso

(96) $$\lambda = f(\mathbf{x})\, dx_1 \wedge \cdots \wedge dx_{r-1} \wedge dx_{r+1} \wedge \cdots \wedge dx_k$$

debido a que cada $(k-1)$-forma es una suma de estos especiales, para $r = 1, \ldots, k$.

Por (85), la frontera del simplex (93) es

$$\partial\sigma = [\mathbf{e}_1, \ldots, \mathbf{e}_k] + \sum_{i=1}^{k} (-1)^i \tau_i$$

en donde

$$\tau_i = [\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \ldots, \mathbf{e}_k]$$

para $i = 1, \ldots, k$. Haciendo

$$\tau_0 = [\mathbf{e}_r, \mathbf{e}_1, \ldots, \mathbf{e}_{r-1}, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_k].$$

Nótese que τ_0 se obtuvo a partir de $[\mathbf{e}_1, \ldots, \mathbf{e}_k]$ por medio de $r-1$ intercambios sucesivos de $\mathbf{e}_r$ y sus vecinos izquierdos. Por esto,

(97) $$\partial\sigma = (-1)^{r-1}\tau_0 + \sum_{i=1}^{k} (-1)^i \tau_i .$$

Cada τ_i tiene a Q^{k-1} como dominio de parámetros.

Si $\mathbf{x} = \tau_0(\mathbf{u})$ y $\mathbf{u} \in Q^{k-1}$, entonces

(98) $$x_j = \begin{cases} u_j & (1 \leq j < r), \\ 1 - (u_1 + \cdots + u_{k-1}) & (j = r), \\ u_{j-1} & (r < j \leq k). \end{cases}$$

Si $1 \leq i \leq k$, $\mathbf{u} \in Q^{k-1}$, y $\mathbf{x} = \tau_i(\mathbf{u})$, entonces

(99) $$x_j = \begin{cases} u_j & (1 \leq j < i), \\ 0 & (j = i), \\ u_{j-1} & (i < j \leq k). \end{cases}$$

Para $0 \leq i \leq k$, sea J_i el Jacobiano del mapeo

(100) $$(u_1, \ldots, u_{k-1}) \to (x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_k)$$

inducido por τ_i. Cuando $i = 0$ e $i = r$, (98) y (99) muestran que (100) es el mapeo identidad. Entonces $J_0 = 1$, $J_r = 1$. Para otra i, el hecho de que $x_i = 0$ en (99) muestra que J_i tiene un renglón de ceros, en consecuencia $J_i = 0$. Con lo anterior,

(101) $$\int_{\tau_i} \lambda = 0 \qquad (i \neq 0, i \neq r),$$

por (35) y (96). Por ende, (97) da

(102) $$\begin{aligned}\int_{\partial\sigma} \lambda &= (-1)^{r-1}\int_{\tau_0} \lambda + (-1)^r \int_{\tau_r} \lambda \\ &= (-1)^{r-1}\int \qquad [f(\tau_0(\mathbf{u})) - f(\tau_r(\mathbf{u}))]\, d\mathbf{u}.\end{aligned}$$

Por otra parte,

$$\begin{aligned}d\lambda &= (D_r f)(\mathbf{x}) dx_r \wedge dx_1 \wedge \cdots \wedge dx_{r-1} \wedge dx_{r+1} \wedge \cdots \wedge dx_k \\ &= (-1)^{r-1}(D_r f)(\mathbf{x})\, dx_1 \wedge \cdots \wedge dx_k\end{aligned}$$

así que

(103) $$\int_{\sigma} d\lambda = (-1)^{r-1}\int_{Q^k} (D_r f)(\mathbf{x})\, d\mathbf{x}.$$

Se evalúa (103) integrando primero con respecto a x_r, sobre el intervalo

$$[0, 1 - (x_1 + \cdots + x_{r-1} + x_{r+1} + \cdots + x_k)],$$

poniendo $(x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_k) = (u_1, \ldots, u_{k-1})$, y viendo con ayuda de (98) que la integral sobre Q^k en (103) es igual a la integral sobre Q^{k-1} en (102). Entonces se cumple (94), y se completa la demostración.

FORMAS CERRADAS Y FORMAS EXACTAS

10.34 Definición Sea ω una k-forma en un conjunto abierto $E \subset R^n$. Si hay una $(k - 1)$-forma λ en E tal que $\omega = d\lambda$, entonces se dice que ω es *exacta en E.*

Si ω es de clase $\mathscr{C}'$ y $d\omega = 0$, se dice entonces que ω es *cerrada.*

El Teorema 10.20(b) muestra que cada forma exacta de clase $\mathscr{C}'$ es cerrada.

En determinados conjuntos E, por ejemplo en los convexos, el inverso es verdadero; es decir los contenidos del Teorema 10.39 (comúnmente conocido como lema de *Poincaré*) y del Teorema 10.40. Los Ejemplos 10.36 y 10.37 exhibirán sin embargo, formas cerradas que no son exactas.

10.35 Observaciones

(*a*) Puede verificarse si una k-forma dada ω es o no cerrada, simplemente diferenciando los coeficientes en la representación estándar de ω. Por ejemplo, la 1-forma

(104) $$\omega = \sum_{i=1}^{n} f_i(\mathbf{x})\, dx_i,$$

en donde $f_i \in \mathscr{C}'(E)$ para algún conjunto abierto $E \subset R^n$, es cerrada si y solo si las ecuaciones

(105) $$(D_j f_i)(\mathbf{x}) = (D_i f_j)(\mathbf{x})$$

se satisfacen para todo i, j en $\{1, \ldots, n\}$ y para todo $\mathbf{x} \in E$.

Nótese que (105) es una condición "puntual"; no involucra propiedades globales que dependen de la forma de E.

Por otro lado, para mostrar que ω es exacta en E, se tiene que probar la existencia de una forma λ, definida en E, y que sea tal que $d\lambda = \omega$. Esto equivale a resolver un sistema de ecuaciones diferenciales parciales, no solo localmente, sino en todo E. Por ejemplo, para mostrar que (104) es exacta en un conjunto E, se tiene que encontrar una función (o 0-forma) $g \in \mathscr{C}'(E)$ tal que

(106) $$(D_i g)(\mathbf{x}) = f_i(\mathbf{x}) \qquad (\mathbf{x} \in E,\ 1 \leq i \leq n).$$

Por supuesto que (105) es una condición necesaria para que pueda resolverse (106).

(*b*) Sea ω una k-forma *exacta* en E. Entonces hay una $(k - 1)$-forma λ en E con $d\lambda = \omega$, y el teorema de Stokes asegura que para cada k-cadena Ψ de clase $\mathscr{C}''$ en E,

(107) $$\int_{\Psi} \omega = \int_{\Psi} d\lambda = \int_{\partial\Psi} \lambda$$

Si Ψ_1 y Ψ_2 son tales cadenas, y tienen las mismas fronteras, se deduce que

$$\int_{\Psi_1} \omega = \int_{\Psi_2} \omega.$$

En particular, *la integral de una k-forma exacta en E es* 0 *sobre cada k-cadena en E cuya frontera es* 0.

Nótese, como un caso especial de esto, que esas integrales de las 1-formas exactas en *E son 0 sobre curvas cerradas (diferenciables) en E.*

(*c*) Sea ω una k-forma *cerrada* en E. Entonces $d\omega = 0$, y el teorema de Stokes asegura que

(108) $$\int_{\partial\Psi} \omega = \int_{\Psi} d\omega = 0$$

para cada $(k + 1)$-cadena Ψ de clase $\mathscr{C}''$ en E.

En otras palabras, *las integrales de las k-formas cerradas en E son* 0 *sobre k-cadenas que sean fronteras de* $(k + 1)$*-cadenas en E.*

(*d*) Sea Ψ una $(k + 1)$-cadena en E y λ una $(k - 1)$-forma en E, ambas de clase $\mathscr{C}''$. Debido a que $d^2\lambda = 0$, dos aplicaciones del teorema de Stokes muestran que

(109) $$\int_{\partial\partial\Psi} \lambda = \int_{\partial\Psi} d\lambda = \int_{\Psi} d^2\lambda = 0.$$

Se concluye que $\partial^2\Psi = 0$. En otras palabras, que *la frontera de una frontera es* 0.

Para una demostración más directa de esto, véase el Ejercicio 16.

10.36 Ejemplo Sea $E = R^2 - \{\mathbf{0}\}$, el plano sin origen. La 1-forma

(110) $$\eta = \frac{x\,dy - y\,dx}{x^2 + y^2}$$

es *cerrada* en $R^2 - \{\mathbf{0}\}$. Esto se verifica fácilmente por diferenciación. Se fija $r > 0$, y se define

(111) $$\gamma(t) = (r \cos t, r \operatorname{sen} t) \qquad (0 \leq t \leq 2\pi).$$

Entonces γ es una curva (un "1-simplex orientado") en $R^2 - \{\mathbf{0}\}$. Como $\gamma(0) = \gamma(2\pi)$, se tiene

(112) $$\partial\gamma = 0.$$

Un cálculo directo muestra que

(113) $$\int_{\gamma} \eta = 2\pi \neq 0.$$

En las Observaciones 10.35(*b*) y (*c*) pueden inferirse de (113) dos conclusiones:

Primera, η *no es exacta en* $R^2 - (\mathbf{0})$, porque de otra forma (112) forzaría a la integral (113) a ser 0.

Segunda, γ *no es la frontera de cualquier* 2*-cadena en* $R^2 - \{\mathbf{0}\}$ (de clase $\mathscr{C}''$), porque de otra forma, el hecho de que η es cerrada forzaría a la integral (113) a ser 0.

10.37 Ejemplo Sea $E = R^3 - \{\mathbf{0}\}$, un 3-espacio sin el origen. Se define

$$\zeta = \frac{x\, dy \wedge dz + y\, dz \wedge dx + z\, dx \wedge dy}{(x^2 + y^2 + z^2)^{3/2}} \tag{114}$$

en donde se ha escrito (x,y,z) en lugar de (x_1,x_2,x_3). La diferenciación muestra que $d\zeta = 0$, así que ζ es una 2-forma cerrada en $R^3 - \{\mathbf{0}\}$.

Sea Σ la 2-cadena en $R^3 - \{\mathbf{0}\}$ que se construyó en el Ejemplo 10.32; recuérdese que Σ es una parametrización de la esfera unitaria en R^3. Si se usa el rectángulo D del Ejemplo 10.32 como dominio de parámetros, se puede calcular fácilmente

$$\int_\Sigma \zeta = \int_D \operatorname{sen} u\, du\, dv = 4\pi \neq 0. \tag{115}$$

Como en el ejemplo anterior, puede concluirse ahora, que ζ *no* es exacta en $R^3 - \{\mathbf{0}\}$ (debido a que como se mostró en el Ejemplo 10.32, $\partial\Sigma = 0$) y que la esfera Σ no es la frontera de cualquier 3-cadena en $R^3 - \{\mathbf{0}\}$ (de clase $\mathscr{C}''$), aunque $\partial\Sigma = 0$.

El siguiente resultado se usará en la demostración del Teorema 10.39.

10.38 Teorema *Supóngase que E es un conjunto abierto convexo en R^n, $f \in \mathscr{C}'(E)$ p es un entero, $1 \leq p \leq n$, y*

$$(D_j f)(\mathbf{x}) = 0 \qquad (p < j \leq n, \mathbf{x} \in E). \tag{116}$$

Entonces existe un $F \in \mathscr{C}'(E)$ tal que

$$(D_p F)(\mathbf{x}) = f(\mathbf{x}), \qquad (D_j F)(\mathbf{x}) = 0 \qquad (p < j \leq n, \mathbf{x} \in E). \tag{117}$$

Demostración Escríbase $\mathbf{x} = (\mathbf{x}', x_p, \mathbf{x}'')$, donde

$$\mathbf{x}' = (x_1, \ldots, x_{p-1}), \ \mathbf{x}'' = (x_{p+1}, \ldots, x_n).$$

(Cuando $p = 1$, $\mathbf{x}'$ no aparece; cuando $p = n$, $\mathbf{x}''$ no aparece.) Sea V el conjunto de todas las $(\mathbf{x}',x_p) \in R^p$ tales que $(\mathbf{x}',x_p,\mathbf{x}'') \in E$ para algún $\mathbf{x}''$. Siendo una proyección de E, V es un conjunto abierto convexo en R^p. Como E es convexo y (116) se cumple, $f(\mathbf{x})$ no depende de $\mathbf{x}''$. Por lo anterior, hay una función φ, con dominio V, tal que

$$f(\mathbf{x}) = \varphi(\mathbf{x}', x_p)$$

para todo $\mathbf{x} \in E$.

Si $p = 1$, V, es un segmento en R^1 (posiblemente no acotado). Se elige $c \in V$ y se define

$$F(\mathbf{x}) = \int_c^{x_1} \varphi(t)\, dt \qquad (\mathbf{x} \in E).$$

Si $p > 1$, sea U el conjunto de todos los $\mathbf{x}' \in R^{p-1}$ tales que $(\mathbf{x}',x_p) \in V$ para algún x_p. Entonces U es un conjunto abierto convexo en R^{p-1}, y hay una función $\alpha \in \mathscr{C}'(U)$ tal que $(\mathbf{x}'\ \alpha(\mathbf{x}')) \in V$ para cada $\mathbf{x}' \in U$; en otras palabras, la gráfica de α está en V (Ejercicio 29). Defínase ahora

$$F(\mathbf{x}) = \int_{\alpha(\mathbf{x}')}^{x_p} \varphi(\mathbf{x}'; t)\, dt \qquad (\mathbf{x} \in E).$$

En cada caso, F satisface (117).

(*Nota*: Recuérdese la convención usual, que $\int_a^b$ significa $-\int_b^a$ si $b < a$.)

10.39 Teorema *Si $E \subset R^n$ es un conjunto abierto convexo, si $k \geq 1$, y si ω es una k-forma de clase $\mathscr{C}'$ en E, y $d\omega = 0$, entonces hay una $(k - 1)$-forma λ en E tal que $\omega = d\lambda$.*

En breve, las formas cerradas son exactas en conjuntos convexos.

Demostración Para $p = 1, \ldots, n$, sea Y_p el conjunto que representa a todas las k-formas ω, de clase $\mathscr{C}'$ en E, cuya presentación estándar

(118) $$\omega = \sum_I f_I(\mathbf{x})\, dx_I$$

no involucra a $dx_{p+1}, \ldots, dx_n$. En otras palabras, $I \subset \{1, \ldots, p\}$, si $f_I(\mathbf{x}) \neq 0$ para algún $\mathbf{x} \in E$.

Se procederá por inducción sobre p.

Supóngase primero que $\omega \in Y_1$. Entonces $\omega = f(\mathbf{x})\, dx_1$. Como $d\omega = 0$, $(D_j f)(\mathbf{x}) = 0$ para $1 < j \leq n$, $\mathbf{x} \in E$. Por el Teorema 10.38 existe un $F \in \mathscr{C}'(E)$ tal que $D_1 F = f$ y $D_j F = 0$ para $1 < j \leq n$. Entonces,

$$dF = (D_1 F)(\mathbf{x})\, dx_1 = f(\mathbf{x})\, dx_1 = \omega.$$

Se toma ahora $p > 1$ y se hacen las siguientes hipótesis de inducción: *Cada k-forma cerrada que pertenece a Y_{p-1} es exacta en E.*

Elíjase $\omega \in Y_p$ de manera que $d\omega = 0$. Por (118) se tiene

(119) $$\sum_I \sum_{j=1}^n (D_j f_I)(\mathbf{x})\, dx_j \wedge dx_I = d\omega = 0.$$

Considere un j fijo, con $p < j \leq n$. Cada I que aparece en (118) está en $\{1, \ldots, p\}$. Si I_1, I_2 son dos de estos k-índices, y si $I_1 \neq I_2$, entonces los $(k + 1)$-índices (I_1, j), (I_2, j) son distintos. Entonces no hay cancelación, y de (119) se concluye que cada coeficiente en (118) satisface

(120) $$(D_j f_I)(\mathbf{x}) = 0 \qquad (\mathbf{x} \in E, p < j \leq n).$$

Se reúnen ahora los términos que en (118) contienen dx_p y se vuelve a escribir ω en la forma

(121) $$\omega = \alpha + \sum_{I_0} f_I(\mathbf{x})\, dx_{I_0} \wedge dx_p,$$

donde $\alpha \in Y_{p-1}$, cada I_0 es un $(k - 1)$-índice creciente en $\{1, \ldots, p - 1\}$, e $I = (I_0, p)$. Por (120), el Teorema 10.38 proporciona funciones $F_I \in \mathscr{C}'(E)$ tales que

(122) $$D_p F_I = f_I, \qquad D_j F_I = 0 \qquad (p < j \leq n).$$

Poniendo

(123) $$\beta = \sum_{I_0} F_I(\mathbf{x})\, dx_{I_0}$$

y definiendo $\gamma = \omega - (-1)^{k-1}\, d\beta$. Ya que β es una $(k - 1)$-forma, se deduce que

$$\begin{aligned}\gamma &= \omega - \sum_{I_0} \sum_{j=1}^{p} (D_j F_I)(\mathbf{x})\, dx_{I_0} \wedge dx_j \\ &= \alpha - \sum_{I_0} \sum_{j=1}^{p-1} (D_j F_I)(\mathbf{x})\, dx_{I_0} \wedge dx_j,\end{aligned}$$

que evidentemente está en Y_{p-1}. Como $d\omega = 0$ y $d^2\beta = 0$, se tiene que $d\gamma = 0$. La hipótesis de inducción muestra por tanto, que $\gamma = d\mu$ para alguna $(k - 1)$-forma μ en E. Si $\lambda = \mu + (-1)^{k-1}\beta$, se concluye que $\omega = d\lambda$.

La demostración se completa por inducción.

10.40 Teorema *Con k, $1 \leq k \leq n$ fijo. Sea $E \subset R^n$ un conjunto abierto en el que cada k-forma cerrada es exacta. Sea T 1-1 un mapeo-$\mathscr{C}''$ de un conjunto E sobre un conjunto abierto $U \subset R^n$ cuyo inverso S es también de clase $\mathscr{C}''$.*

Entonces cada k-forma cerrada en U es exacta en U.

Nótese que cada conjunto E abierto convexo satisface la hipótesis presente, debido al Teorema 10.39. La relación entre E y U puede expresarse diciendo que son *$\mathscr{C}''$-equivalentes.*

Entonces cada forma cerrada es exacta en cualquier conjunto que sea $\mathscr{C}''$ -equivalente a un conjunto convexo abierto.

Demostración Sea ω una k-forma en U, con $d\omega = 0$. Por el Teorema 10.22(c), ω_T es una k-forma en E para la cual $d(\omega_T) = 0$. En consecuencia, $\omega_T = d\lambda$ para alguna $(k - 1)$-forma λ en E. Por el Teorema 10.23 y otra aplicación del Teorema 10.22(c),

$$\omega = (\omega_T)_S = (d\lambda)_S = d(\lambda_S).$$

Debido a que λ_s es una $(k - 1)$-forma en U, ω es exacta en U.

10.41 Observación En aplicaciones, las celdas (véase Definición 2.17) son dominios de parámetros más convenientes que los simplex. Si todo el desarrollo se hubiese basado en celdas en lugar de hacerlo en simplex, hubiese sido incluso más simple el cálculo que aparece en la demostración del teorema de Stokes. (De esta forma se hace en el libro de Spivak). La razón de haber preferido simplex es que la definición de la frontera de un simplex orientado parece ser más fácil y más natural que la correspondiente para una celda. (Véase el Ejercicio 19.) También la partición de conjuntos en simplex (denominada "triangulación") desempeña un papel importante en topología, y existen conexiones poderosas entre algunos aspectos de la topología, por un lado, y las formas diferenciales por el otro. Estas se sugirieron en la sección 10.35. Una buena introducción a este tema se encuentra en el libro de Singer y Thorpe.

Debido a que cada celda puede triangularse, se puede considerar como una cadena. Esto se hizo en el Ejemplo 10.32 para el caso de dimensión 2; para el de dimensión 3, véase el Ejercicio 18.

El lema de Poincaré (Teorema 10.39) se puede probar de diversas formas. Por ejemplo, véase la página 94 del libro de Spivak, o la 280 del de Fleming. En los Ejercicios 24 y 27 se indican dos demostraciones de algunos casos especiales.

ANÁLISIS VECTORIAL

Se concluirá este capítulo con algunas aplicaciones de los temas anteriores a los teoremas referentes al análisis vectorial en R^3. Estos son casos especiales de teoremas sobre formas diferenciales, pero se establecen comúnmente con diferente terminología. Nos enfrentamos a la tarea de traducir de un lenguaje a otro.

10.42 Campos vectoriales Sea $\mathbf{F} = F_1\mathbf{e}_1 + F_2\mathbf{e}_2 + F_3\mathbf{e}_3$ un mapeo continuo de un conjunto abierto $E \subset R^3$ en R^3. A $\mathbf{F}$ se le llama algunas veces campo vectorial, especialmente en Física, porque a cada punto de E, $\mathbf{F}$ asocia un vector. Con cada $\mathbf{F}$ tal, se asocia una 1-forma

(124) $$\lambda_{\mathbf{F}} = F_1\,dx + F_2\,dy + F_3\,dz$$

y una 2-forma

(125) $$\omega_{\mathbf{F}} = F_1\,dy \wedge dz + F_2\,dz \wedge dx + F_3\,dx \wedge dy.$$

Se usará aquí, y en el resto del capítulo, la acostumbrada notación (x,y,z) en vez de (x_1,x_2,x_3).

Inversamente, es claro que cada 1-forma λ en E es $\lambda_{\mathbf{F}}$ para algún campo vectorial $\mathbf{F}$ en E, y que cada 2-forma ω es $\omega_{\mathbf{F}}$ para algún $\mathbf{F}$. En R^3, el estudio de 1-formas y 2-formas es entonces coextensivo con el estudio de campos vectoriales.

Si $u \in \mathscr{C}'(E)$ es una función real, entonces su *gradiente*

$$\nabla u = (D_1u)\mathbf{e}_1 + (D_2u)\mathbf{e}_2 + (D_3u)\mathbf{e}_3$$

es un ejemplo de un campo vectorial en E.

Supóngase ahora que $\mathbf{F}$ es un campo vectorial en E, de clase $\mathscr{C}'$. Su *rotacional* $\nabla \times \mathbf{F}$ es el campo vectorial definido en E por

$$\nabla \times \mathbf{F} = (D_2F_3 - D_3F_2)\mathbf{e}_1 + (D_3F_1 - D_1F_3)\mathbf{e}_2 + (D_1F_2 - D_2F_1)\mathbf{e}_3$$

y su *divergencia* es la función real $\nabla \cdot \mathbf{F}$ definida en E por

$$\nabla \cdot \mathbf{F} = D_1F_1 + D_2F_2 + D_3F_3.$$

Estas cantidades tienen varias interpretaciones físicas. Para mayores detalles véase el libro de O. D. Kellogg.

A continuación se tienen algunas relaciones entre el gradiente, el rotacional y la divergencia.

10.43 Teorema *Supóngase que E es un conjunto abierto en R^3, $u \in \mathscr{C}''(E)$ y $\mathbf{G}$ es un campo vectorial en E, de clase C''.*

(*a*) *Si $\mathbf{F} = \nabla u$, entonces $\nabla \times \mathbf{F} = \mathbf{0}$.*
(*b*) *Si $\mathbf{F} = \nabla \times \mathbf{G}$, entonces $\nabla \cdot \mathbf{F} = 0$.*

Además, si E es $\mathscr{C}''$-equivalente a un conjunto convexo, entonces (*a*) *y* (*b*) *tienen sus inversos, en los cuales se supone que $\mathbf{F}$ es un campo vectorial en E, de clase $\mathscr{C}'$:*

(*a'*) *Si $\nabla \times \mathbf{F} = \mathbf{0}$, entonces $\mathbf{F} = \nabla u$ para algún $u \in \mathscr{C}''(E)$.*
(*b'*) *Si $\nabla \cdot \mathbf{F} = 0$, entonces $\mathbf{F} = \nabla \times \mathbf{G}$ para algún campo vectorial $\mathbf{G}$ en E, de clase $\mathscr{C}''$.*

Demostración Si se comparan las definiciones de ∇u, $\nabla \times \mathbf{F}$, y $\nabla \cdot \mathbf{F}$ con las formas diferenciales $\lambda_{\mathbf{F}}$ y $\omega_{\mathbf{F}}$ dadas por (124) y (125), se obtienen las siguientes cuatro proposiciones:

$$\begin{array}{rll} \mathbf{F} = \nabla u & \text{si y solo si} & \quad = du. \\ \nabla \times \mathbf{F} = \mathbf{0} & \text{si y solo si} & d\lambda_{\mathbf{F}} = 0. \\ \mathbf{F} = \nabla \times \mathbf{G} & \text{si y solo si} & \omega_{\mathbf{F}} = d\lambda_{\mathbf{G}}. \\ \nabla \cdot \mathbf{F} = 0 & \text{si y solo si} & d\omega_{\mathbf{F}} = 0. \end{array}$$

Si ahora $\mathbf{F} = \nabla u$, entonces $\lambda_{\mathbf{F}} = du$, por eso $d\lambda_{\mathbf{F}} = d^2u = 0$ (Teorema 10.20), lo que significa que $\nabla \times \mathbf{F} = \mathbf{0}$. Y por lo anterior ($a$) queda demostrado.

Por lo que a (a') se refiere, la hipótesis se amplía diciendo que $d\lambda_{\mathbf{F}} = 0$ en E. Por el Teorema 10.40, $\lambda_{\mathbf{F}} = du$ para alguna 0-forma u. Por consiguiente, $\mathbf{F} = \nabla u$.

Las demostraciones de (b) y (b') siguen los mismos pasos.

10.44 Elementos de volumen La k-forma

$$dx_1 \wedge \cdots \wedge dx_k$$

se denomina el elemento de volumen en R^k. Se representa normalmente por dV (o por dV_k en el caso en que se necesite indicar explícitamente la dimensión), y se usa la notación

$$\int_{\Phi} f(\mathbf{x})\, dx_1 \wedge \cdots \wedge dx_k = \int_{\Phi} f\, dV \tag{126}$$

cuando Φ es una k-superficie orientada positivamente en R^k y f es una función continua sobre el rango de Φ.

La razón de usar esta terminología es muy sencilla: Si D es un dominio de parámetros en R^k, y si Φ es un mapeo- $\mathscr{C}'$ 1-1 de D en R^k, con Jacobiano J_Φ positivo, entonces el miembro izquierdo de (126) es, por (35) y el Teorema 10.9,

$$\int_D f(\Phi(\mathbf{u}))J_\Phi(\mathbf{u})\, d\mathbf{u} = \int_{\Phi(D)} f(\mathbf{x})\, d\mathbf{x},$$

En particular, cuando $f = 1$, (126) define el *volumen* de Φ. Ya se vio un caso especial de esto en (36).

La notación común para dV_2 es dA.

10.45 Teorema de Green *Supóngase que E es un conjunto abierto en R^2, $\alpha \in \mathscr{C}'(E)$, $\beta \in \mathscr{C}'(E)$ y Ω es un subconjunto cerrado de E, con frontera $\partial\Omega$ orientada positivamente, como se describió en la sección* 10.31. *Entonces,*

(127) $$\int_{\partial\Omega}(\alpha\,dx+\beta\,dy)=\int_{\Omega}\left(\frac{\partial\beta}{\partial x}-\frac{\partial\alpha}{\partial y}\right)dA.$$

Demostración Poniendo $\lambda = \alpha\,dx + \beta\,dy$. Entonces

$$\begin{aligned} d\lambda &= (D_2\alpha)\,dy \wedge dx + (D_1\beta)\,dx \wedge dy \\ &= (D_1\beta - D_2\alpha)\,dA, \end{aligned}$$

y (127) es lo mismo que

$$\int_{\partial\Omega}\lambda=\int_{\Omega}d\lambda,$$

que por el Teorema 10.33 es verdadera.

Con $\alpha(x,y) = -y$ y $\beta(x,y) = x$, (127) se vuelve

(128) $$\tfrac{1}{2}\int_{\partial\Omega}(x\,dy-y\,dx)=A(\Omega),$$

que es el área de Ω.

Con $\alpha = 0$, $\beta = x$, se obtiene una fórmula similar. El Ejemplo 10.12(*b*) contiene un caso especial de esto.

10.46 Elementos de área en R^3. Sea Φ una 2-superficie en R^3, de clase $\mathscr{C}'$ con dominio de parámetros $D \subset R^2$. Asóciese a cada punto $(u,v) \in D$ el vector

(129) $$\mathbf{N}(u,v)=\frac{\partial(y,z)}{\partial(u,v)}\mathbf{e}_1+\frac{\partial(z,x)}{\partial(u,v)}\mathbf{e}_2+\frac{\partial(x,y)}{\partial(u,v)}\mathbf{e}_3\,.$$

Los Jacobianos en (129) corresponden a la ecuación

(130) $$(x,y,z)=\Phi(u,v).$$

Si f es una función continua sobre $\Phi(D)$, la *integral de área* de f sobre Φ se define como

(131) $$\int_{\Phi}f\,dA=\int_{D}f(\Phi(u,v))\,|\mathbf{N}(u,v)|\,du\,dv.$$

En particular, cuando $f = 1$ se obtiene el *área* de Φ, es decir,

(132) $$A(\Phi)=\int_{D}|\mathbf{N}(u,v)|\,du\,dv.$$

En lo que sigue se mostrará que (131) y su caso especial, (132), son definiciones razonables. También se describirán las características del vector **N**.

Se escribe $\Phi = \varphi_1\mathbf{e}_1 + \varphi_2\mathbf{e}_2 + \varphi_3\mathbf{e}_3$, y se fija un punto $\mathbf{p}_0 = (u_0, v_0) \in D$, si se pone $\mathbf{N} = \mathbf{N}(\mathbf{p}_0)$, y

$$\alpha_i = (D_1\varphi_i)(\mathbf{p}_0), \qquad \beta_i = (D_2\varphi_i)(\mathbf{p}_0) \qquad (i = 1, 2, 3) \tag{133}$$

ahora, sea $T \in L(R^2, R^3)$ la transformación lineal dada por

$$T(u, v) = \sum_{i=1}^{3} (\alpha_i u + \beta_i v)\mathbf{e}_i. \tag{134}$$

Nótese que de acuerdo con la Definición 9.11, $T = \Phi'(\mathbf{p}_0)$.

Admitamos ahora que el rango de T es 2. (Si fuera 1 ó 0, entonces $\mathbf{N} = \mathbf{0}$, y el plano tangente mencionado a continuación degenera en una línea o en un punto.) El rango del mapeo afín

$$(u, v) \to \Phi(\mathbf{p}_0) + T(u, v)$$

es entonces un plano Π, llamado *plano tangente* a Φ en $\mathbf{p}_0$. [Sería preferible llamarle Π al plano tangente en $\Phi(\mathbf{p}_0)$, en vez de en $\mathbf{p}_0$; si Φ no es uno-a-uno, esto conduce a dificultades.]

Si en (129) se usa (133), obtenemos

$$\mathbf{N} = (\alpha_2\beta_3 - \alpha_3\beta_2)\mathbf{e}_1 + (\alpha_3\beta_1 - \alpha_1\beta_3)\mathbf{e}_2 + (\alpha_1\beta_2 - \alpha_2\beta_1)\mathbf{e}_3, \tag{135}$$

y (134) muestra que

$$T\mathbf{e}_1 = \sum_{i=1}^{3} \alpha_i\mathbf{e}_i, \qquad T\mathbf{e}_2 = \sum_{i=1}^{3} \beta_i\mathbf{e}_i. \tag{136}$$

Ahora un cálculo sencillo conduce a

$$\mathbf{N} \cdot (T\mathbf{e}_1) = 0 = \mathbf{N} \cdot (T\mathbf{e}_2). \tag{137}$$

Por lo anterior, **N** es perpendicular a Π. Por tanto se le llama la *normal a Φ en $\mathbf{p}_0$*.

Una segunda propiedad de **N**, que también se verifica mediante un cálculo directo basado en (135) y (136), es que el determinante de la transformación lineal de R^3 que lleva $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ en $\{T\mathbf{e}_1, T\mathbf{e}_2, \mathbf{N}\}$ es $|\mathbf{N}|^2 > 0$ (véase Ejercicio 30). El 3-simplex

$$[\mathbf{0}, T\mathbf{e}_1, T\mathbf{e}_2, \mathbf{N}] \tag{138}$$

es entonces *orientado positivamente*.

La tercera propiedad de $\mathbf{N}$ que se usará, es consecuencia de las dos primeras: El determinante mencionado anteriormente, cuyo valor es $|\mathbf{N}|^2$, es el volumen del paralelepípedo con lados $[\mathbf{0}, T\mathbf{e}_1]$, $[\mathbf{0}, T\mathbf{e}_2]$, $[\mathbf{0}, \mathbf{N}]$. Por (137), $[\mathbf{0}, \mathbf{N}]$ es perpendicular a los otros dos lados. *El área del paralelogramo con vértices*

$$\mathbf{0}, T\mathbf{e}_1, T\mathbf{e}_2, T(\mathbf{e}_1 + \mathbf{e}_2) \tag{139}$$

es por tanto $|\mathbf{N}|$.

Este paralelogramo es la imagen bajo T del cuadrado unitario en R^2. Si E es cualquier rectángulo en R^2, se deduce (por la linealidad de T) que el área del paralelogramo $T(E)$ es

$$A(T(E)) = |\mathbf{N}|A(E) = \int_E |\mathbf{N}(u_0, v_0)|\, du\, dv. \tag{140}$$

Se concluye que (132) es correcta cuando Φ es afín. Para justificar la definición (132) en el caso general, divídase en pequeños rectángulos D, elíjase un punto (u_0, v_0) en cada uno y reemplácese en cada rectángulo Φ por el plano tangente correspondiente. La suma de las áreas de los paralelogramos resultantes, obtenida vía (140), es entonces una aproximación a $A(\Phi)$. Finalmente, se puede justificar (131) a partir de (132) aproximando f por medio de funciones escalonadas.

10.47 Ejemplo Sea $0 < a < b$ fijo. Sea K la 3-celda determinada por

$$0 \le t \le a, \qquad 0 \le u \le 2\pi, \qquad 0 \le v \le 2\pi.$$

Las ecuaciones

$$\begin{aligned} x &= t\cos u \\ y &= (b + t\,\mathrm{sen}\, u)\cos v \\ z &= (b + t\,\mathrm{sen}\, u)\,\mathrm{sen}\, v \end{aligned} \tag{141}$$

describen un mapeo Ψ de R^3 en R^3 que es 1-1 en el interior de K, tal que $\Psi(K)$ es un toroide sólido. Su Jacobiano es

$$J_\Psi = \frac{\partial(x, y, z)}{\partial(t, u, v)} = t(b + t\,\mathrm{sen}\, u)$$

y es positivo sobre K, excepto sobre la cara $t = 0$. Si se integra J_Ψ sobre K, se obtiene

$$\mathrm{vol}\,(\Psi(K)) = 2\pi^2 a^2 b$$

que es el volumen del toroide sólido.

Considérese ahora la 2-cadena $\Phi = \partial\Psi$. (Véase Ejercicio 19.) Ψ mapea las caras $u = 0$ y $u = 2\pi$ de K sobre la misma franja cilíndrica, pero con orientaciones opuestas. Ψ mapea las caras $v = 0$ y $v = 2\pi$ sobre el mismo disco circular, pero con orientaciones opuestas. Ψ mapea la cara $t = 0$ sobre el círculo, que contribuye 0 a la 2-cadena $\partial\Psi$. (Los Jacobianos relevantes son 0.) Por esto Φ es simplemente la 2-superficie obtenida haciendo $t = a$ en (141), con dominio de parámetros D, que es el cuadrado definido por $0 \leq u \leq 2\pi$, $0 \leq v \leq 2\pi$.

De acuerdo con (129) y (141), la normal a Φ en $(u, v) \in D$ es por eso el vector

$$\mathbf{N}(u, v) = a(b + a\,\mathrm{sen}\,u)\mathbf{n}(u, v)$$

en donde

$$\mathbf{n}(u, v) = (\cos u)\mathbf{e}_1 + (\mathrm{sen}\,u \cos v)\mathbf{e}_2 + (\mathrm{sen}\,u \sin v)\mathbf{e}_3 .$$

Ya que $|\mathbf{n}(u, v)| = 1$, se tiene que $|\mathbf{N}(u, v)| = a(b + a \,\mathrm{sen}\, u)$, integrando esto sobre D, (131) nos da

$$A(\Phi) = 4\pi^2 ab$$

lo cual es el área de la superficie del toroide.

Si se piensa que $\mathbf{N} = \mathbf{N}(u, v)$ es un segmento de recta dirigido, apuntando de $\Phi\,(u, v)$ hacia $\Phi(u, v) + \mathbf{N}(u, v)$, entonces $\mathbf{N}$ apunta *hacia afuera*, es decir hacia afuera de $\Psi(K)$. Esto es por que $J_\Psi > 0$ cuando $t = a$.

Tómese por ejemplo $u = v = \pi/2$, $t = a$. Esto da el valor de z más grande sobre $\Psi(K)$, y $\mathbf{N} = a(b + a)\mathbf{e}_3$ apunta "hacia arriba" para esta elección de (u, v).

10.48 Integrales de 1-formas en R^3 Sea γ una curva- $\mathscr{C}'$ en un conjunto abierto $E \subset R^3$ con intervalo de parámetros $[0, 1]$, sea un campo vectorial en E, como en la sección 10.42, y defínase $\lambda_{\mathbf{F}}$ por (124). La integral de $\lambda_{\mathbf{F}}$ sobre γ puede volverse a escribir de la siguiente manera.

Para cualquier $u \in [0, 1]$,

$$\gamma'(u) = \gamma_1'(u)\mathbf{e}_1 + \gamma_2'(u)\mathbf{e}_2 + \gamma_3'(u)\mathbf{e}_3$$

se le llama el *vector tangente* a γ en u. Se define $\mathbf{t} = \mathbf{t}(u)$ como el vector unitario en la dirección de $\gamma'(u)$. Por eso

$$\gamma'(u) = |\gamma'(u)|\,\mathbf{t}(u).$$

[Si $\gamma'(u) = \mathbf{0}$ para algún u, póngase $\mathbf{t}(u) = \mathbf{e}_1$; cualquier otra elección también deberá hacerse así.] Por (35) se tiene

(142)
$$\begin{aligned}\int_\gamma \lambda_{\mathbf{F}} &= \sum_{i=1}^{3} \int_0^1 F_i(\gamma(u))\gamma_i'(u)\,du \\ &= \int_0^1 \mathbf{F}(\gamma(u)) \cdot \gamma'(u)\,du \\ &= \int_0^1 \mathbf{F}(\gamma(u)) \cdot \mathbf{t}(u)\,|\gamma'(u)|\,du.\end{aligned}$$

Por el Teorema 6.27, es razonable llamar a $|\gamma'(u)|\,du$ *elemento de longitud de arco a lo largo de* γ. La notación acostumbrada para éste es ds, y (142) puede volverse a escribir en la forma

(143)
$$\int_\gamma \lambda_{\mathbf{F}} = \int_\gamma (\mathbf{F} \cdot \mathbf{t})\,ds.$$

Como $\mathbf{t}$ es un vector unitario tangente a γ, $\mathbf{F} \cdot \mathbf{t}$ se llama la *componente tangencial* de $\mathbf{F}$ a lo largo de γ.

El miembro derecho de (143) debe considerarse como una abreviación de la última integral de (142). La clave aquí es que $\mathbf{F}$ está definido sobre el rango de γ, mientras que $\mathbf{t}$ está definido sobre $[0, 1]$; por eso $\mathbf{F} \cdot \mathbf{t}$ tiene que interpretarse adecuadamente. Cuando γ es uno-a-uno, entonces $\mathbf{t}(u)$ puede reemplazarse, por supuesto, con $\mathbf{t}(\gamma(u))$, y esta dificultad desaparece.

10.49 Integrales de 2-formas en R^3 Sea Φ una 2-superficie en un conjunto abierto $E \subset R^3$, de clase $\mathscr{C}'$, con dominio de parámetros $D \subset R^2$. Sea $\mathbf{F}$ un campo vectorial en E, y $\omega_{\mathbf{F}}$ definida por (125). Como en la sección anterior, se obtendrá una representación diferente de la integral de $\omega_{\mathbf{F}}$ sobre Φ.

Por (35) y (129), se tiene

$$\begin{aligned}\int_\Phi \omega_{\mathbf{F}} &= \int_\Phi (F_1\,dy \wedge dz + F_2\,dz \wedge dx + F_3\,dx \wedge dy) \\ &= \int_D \left\{(F_1 \circ \Phi)\frac{\partial(y, z)}{\partial(u, v)} + (F_2 \circ \Phi)\frac{\partial(z, x)}{\partial(u, v)} + (F_3 \circ \Phi)\frac{\partial(x, y)}{\partial(u, v)}\right\} du\,dv \\ &= \int_D \mathbf{F}(\Phi(u, v)) \cdot \mathbf{N}(u, v)\,du\,dv.\end{aligned}$$

Sea ahora $\mathbf{n} = \mathbf{n}(u,v)$ el vector unitario en la dirección de $\mathbf{N}(u,v)$. [Si $\mathbf{N}(u,v) = \mathbf{0}$ para algún $(u,v) \in D$, se toma $\mathbf{n}(u,v) = \mathbf{e}_1$.] Entonces $\mathbf{N} = |\mathbf{N}|\mathbf{n}$, y por tanto, la última integral se convierte en

$$\int_D \mathbf{F}(\Phi(u, v)) \cdot \mathbf{n}(u, v)\,|\mathbf{N}(u, v)|\,du\,dv.$$

Por (131) se puede escribir finalmente esto como

$$\int_{\Phi} \omega_{\mathbf{F}} = \int_{\Phi} (\mathbf{F} \cdot \mathbf{n})\, dA. \tag{144}$$

Aquí también se aplica la observaciòn que se hizo al final de la sección 10.48, tocante al significado de $\mathbf{F} \cdot \mathbf{n}$.

Ahora ya puede establecerse el teorema de Stokes en su forma original.

10.50 Fórmula de Stokes *Si* $\mathbf{F}$ *es un campo-vectorial de clase* $\mathscr{C}'$ *en un conjunto abierto* $E \subset R^3$, *y si* Φ *es una* 2*-superficie de clase* $\mathscr{C}''$ *en* E, *entonces*

$$\int_{\Phi} (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dA = \int_{\partial\Phi} (\mathbf{F} \cdot \mathbf{t})\, ds. \tag{145}$$

Demostración Si se hace $\mathbf{H} = \nabla \times \mathbf{F}$, entonces, como en la demostración del Teorema 10.43, se tiene

$$\omega_{\mathbf{H}} = d\lambda_{\mathbf{F}}. \tag{146}$$

De esto

$$\int_{\Phi} (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dA = \int_{\Phi} (\mathbf{H} \cdot \mathbf{n})\, dA = \int_{\Phi} \omega_{\mathbf{H}}$$
$$= \int_{\Phi} d\lambda_{\mathbf{F}} = \int_{\partial\Phi} \lambda_{\mathbf{F}} = \int_{\partial\Phi} (\mathbf{F} \cdot \mathbf{t})\, ds.$$

Aquí se usó la definición de $\mathbf{H}$, después (144) con $\mathbf{H}$ en lugar de $\mathbf{F}$; luego (146), y después, lo que se puede considerar el paso principal, el Teorema 10.33, y, finalmente (143), ampliada en la forma obvia, de curvas a 1-cadenas.

10.51 El teorema de la divergencia *Si* $\mathbf{F}$ *es un campo vectorial de clase* $\mathscr{C}'$ *en un conjunto abierto* $E \subset R^3$, *y si* Ω *es un subconjunto cerrado de* E *con frontera orientada positivamente* $\partial\Omega$ *(como se describió en la Sec.* 10.31), *entonces*

$$\int_{\Omega} (\nabla \cdot \mathbf{F})\, dV = \int_{\partial\Omega} (\mathbf{F} \cdot \mathbf{n})\, dA. \tag{147}$$

Demostración Por (125) se tiene

$$d\omega_{\mathbf{F}} = (\nabla \cdot \mathbf{F})\, dx \wedge dy \wedge dz = (\nabla \cdot \mathbf{F})\, dV.$$

De ahí que, aplicando el Teorema 10.33 a la 2-forma $\omega_{\mathbf{F}}$ y (144),

$$\int_{\Omega} (\nabla \cdot \mathbf{F})\, dV = \int_{\Omega} d\omega_{\mathbf{F}} = \int_{\partial\Omega} \omega_{\mathbf{F}} = \int_{\partial\Omega} (\mathbf{F} \cdot \mathbf{n})\, dA,$$

EJERCICIOS

1. Sea H un conjunto compacto convexo en R^k, con interior no vacío. Sea $f \in \mathscr{C}'(H)$, haciendo $f(\mathbf{x}) = 0$ en el complemento de H, y definiendo $\int_H f$ como se hizo en la Definición 10.3.

   Demostrar que $\int_H f$ es independiente del orden en el cual se efectúen las k integraciones.

   Sugerencia: Aproxímese f por funciones que sean continuas sobre R^k y cuyos soportes estén en H, como se hizo en el Ejemplo 10.4.

2. Si $i = 1,2,3,\ldots$, sea $\varphi_i \in \mathscr{C}(R^1)$ con soporte en $(2^{-i}, 2^{1-i})$, tal que $\int \varphi_i = 1$. Se hace

$$f(x, y) = \sum_{i=1}^{\infty} [\varphi_i(x) - \varphi_{i+1}(x)]\varphi_i(y)$$

   Entonces f tiene un soporte compacto en R^2, f es continua excepto en (0, 0), y

$$\int dy \int f(x, y)\, dx = 0 \qquad \text{pero} \qquad \int dx \int f(x, y)\, dy = 1.$$

   Observar que f no es acotada en cada vecindad de (0, 0).

3. (*a*) Si F está dado como en el Teorema 10.7, y se pone $\mathbf{A} = \mathbf{F}'(0)$, $\mathbf{F}_1(\mathbf{x}) = \mathbf{A}^{-1}\mathbf{F}(\mathbf{x})$, entonces $\mathbf{F}_1'(\mathbf{0}) = I$. Mostrar que en alguna vecindad de $\mathbf{0}$,

$$\mathbf{F}_1(\mathbf{x}) = \mathbf{G}_n \circ \mathbf{G}_{n-1} \circ \cdots \circ \mathbf{G}_1(\mathbf{x})$$

   para mapeos primitivos $\mathbf{G}_1, \ldots, \mathbf{G}_n$. Esto da otra versión del Teorema 10.7:

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}'(\mathbf{0})\mathbf{G}_n \circ \mathbf{G}_{n-1} \circ \cdots \circ \mathbf{G}_1(\mathbf{x}).$$

   (*b*) Demostrar que el mapeo $(x,y) \to (y,x)$ de R^2 sobre R^2 no es la composición de dos mapeos primitivos cualesquiera, en cualquier vecindad del origen. (Esto muestra que el flip B_i no se puede omitir en el postulado del Teorema 10.7.)

4. Si para $(x,y) \in R^2$ se define

$$\mathbf{F}(x, y) = (e^x \cos y - 1, e^x \operatorname{sen} y).$$

   Demostrar que $\mathbf{F} = \mathbf{G}_2 \circ \mathbf{G}_1$, donde

$$\mathbf{G}_1(x, y) = (e^x \cos y - 1, y)$$

$$\mathbf{G}_2(u, v) = (u, (1 + u) \tan v)$$

   son primitivos en alguna vecindad de (0, 0).

   Calcular los Jacobianos de $\mathbf{G}_1$, $\mathbf{G}_2$, $\mathbf{F}$ en (0, 0). Definir

$$\mathbf{H}_2(x, y) = (x, e^x \operatorname{sen} y)$$

y determinar

$$\mathbf{H}_1(u, v) = (h(u, v), v)$$

de manera que $\mathbf{F} = \mathbf{H}_1 \circ \mathbf{H}_2$ sea alguna vecindad de (0, 0).

5. Formular y demostrar un teorema análogo al 10.8, en el cual K sea un subconjunto compacto de un espacio métrico arbitrario. (Reemplazar las funciones φ_i que aparecen en la demostración del Teorema 10.8 por funciones del tipo construido en el Ejercicio 22 del Cap. 4.)

6. Reforzar la conclusión del Teorema 10.8 mostrando que las funciones ψ_i pueden hacerse diferenciables, e incluso infinitamente diferenciables. (Al construir las funciones auxiliares φ_i usar el Ejercicio 1 del Cap. 8.)

7. (*a*) Mostrar que el simplex Q^k es el subconjunto convexo de R^k más pequeño, que contiene a $\mathbf{0}$, $\mathbf{e}_1, \ldots, \mathbf{e}_k$.

(*b*) Mostrar que los mapeos afines llevan conjuntos convexos en conjuntos convexos.

8. Sea H el paralelogramo de R^2 cuyos vértices son (1, 1), (3, 2), (4, 5), (2, 4). Encontrar el mapeo afín T que manda (0, 0) a (1, 1), (1, 0) a (3, 2), (0, 1) a (2, 4). Mostrar que $J_T = 5$. Después, usar T para convertir la integral

$$\alpha = \int_H e^{x-y}\, dx\, dy$$

en una integral sobre I^2 y entonces calcular α.

9. Se define $(x, y) = T(r, \theta)$ sobre el rectángulo

$$0 \le r \le a, \qquad 0 \le \theta \le 2\pi$$

por medio de las ecuaciones

$$x = r \cos \theta, \qquad y = r \operatorname{sen} \theta.$$

Mostrar que T mapea este rectángulo sobre el disco cerrado D con centro en (0, 0) y rado a, que T es 1-1 en el interior del rectángulo, y que $J_T(r, \theta) = r$. Si $f \in \mathscr{C}(D)$, demostrar la fórmula de integración en coordenadas polares:

$$\int_D f(x, y)\, dx\, dy = \int_0^a \int_0^{2\pi} f(T(r, \theta))r\, dr\, d\theta.$$

Sugerencia: Sea D_0 el interior de D, menos el intervalo de (0, 0) a (0, a). Como se estableció, el Teorema 10.9 se aplica a funciones continuas f cuyo soporte está en D_0. Para suprimir esta restricción, procédase como en el Ejemplo 10.4.

10. Si $a \to \infty$ en el Ejercicio 9, demostrar que

$$\int_{R^2} f(x, y)\, dx\, dy = \int_0^\infty \int_0^{2\pi} f(T(r, \theta))r\, dr\, d\theta,$$

para funciones continuas f que decrecen suficientemente rápido cuando $|x| + |y| \to \infty$. (Encontrar una formulación más precisa.) Aplicar esto a

$$f(x, y) = \exp(-x^2 - y^2)$$

para deducir la fórmula (101) del capítulo 8.

11. Definir $(u,v) = T(s,t)$ sobre la franja

$$0 < s < \infty, \qquad 0 < t < 1$$

haciendo $u = s - st$, $v = s\,t$. Mostrar que T es un mapeo 1-1 de la franja sobre el cuadrante positivo Q en R^2. Mostrar que $J_T(s,t) = s$.

Para $x > 0$, $y > 0$, integrar

$$u^{x-1}e^{-u}v^{y-1}e^{-v}$$

sobre Q, usar el Teorema 10.9 para convertir la integral a una sobre la franja, y deducir la fórmula (96) del capítulo 8, de esta manera.

(Para esta aplicación, tiene que ampliarse el Teorema 10.9 de tal manera que cubra determinadas integrales impropias. Proporcionar esta ampliación.)

12. Sea I^k el conjunto de todos los ($\mathbf{u} = u_1, \ldots, u_k) \in R^k$ con $0 \leq u_i \leq 1$ para todo i; sea Q^k el conjunto de todos los $\mathbf{x} = (x_1, \ldots, x_k) \in R^k$ con $x_i \geq 0$, $\Sigma x_i \leq 1$. (I^k es el cubo unitario; Q^k es el simplex estándar en R^k.) Si se define $\mathbf{x} = T(\mathbf{u})$ por

$$\begin{aligned} x_1 &= u_1 \\ x_2 &= (1 - u_1)u_2 \\ &\cdots\cdots\cdots\cdots\cdots \\ x_k &= (1 - u_1)\cdots(1 - u_{k-1})u_k. \end{aligned}$$

Mostrar que

$$\sum_{i=1}^{k} x_i = 1 - \prod_{i=1}^{k}(1 - u_i).$$

Mostrar que T mapea I^k sobre Q^k, que T es 1-1 en el interior de I^k, y que su inverso S está definido en el interior de Q^k por $u_1 = x_1$ y

$$u_i = \frac{x_i}{1 - x_1 - \cdots - x_{i-1}}$$

para $i = 2, \ldots, k$. Mostrar que

$$J_T(\mathbf{u}) = (1 - u_1)^{k-1}(1 - u_2)^{k-2}\cdots(1 - u_{k-1}),$$

y

$$J_S(\mathbf{x}) = [(1 - x_1)(1 - x_1 - x_2)\cdots(1 - x_1 - \cdots - x_{k-1})]^{-1}.$$

13. Si $r_1, \ldots, r_k$ son enteros no negativos, demostrar que

$$\int_{Q^k} x_1^{r_1} \cdots x_k^{r_k}\, dx = \frac{r_1! \cdots r_k!}{(k + r_1 + \cdots + r_k)!}$$

Sugerencia: Usar el Ejercicio 12 y los Teoremas 10.9 y 8.20.

Nótese que el caso especial $r_1 = \cdots = r_k = 0$ muestra que el volumen de Q^k es $1/k!$.

14. Demostrar la fórmula (46).

15. Si ω y λ son k- y m-formas, respectivamente, demostrar que

$$\omega \wedge \lambda = (-1)^{km} \lambda \wedge \omega.$$

16. Si $k \geq 2$ y $\sigma = [\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_k]$ es un k-simplex afín orientado, demostrar que $\partial^2\sigma = 0$, directamente de la definición del operador frontera ∂. A partir de esto, deducir que $\partial^2\Psi = 0$ para cada cadena Ψ.

Sugerencia: Hacerlo primero para $k = 2$, $k = 3$. En general, si $i < j$, sea σ_{ij} el $(k - 2)$-simplex obtenido suprimiendo $\mathbf{p}_i$ y $\mathbf{p}_j$ de σ. Mostrar que cada σ_{ij} aparece dos veces en $\partial^2\sigma$, con signo opuesto.

17. Si se hace $J^2 = \tau_1 + \tau_2$, donde

$$\tau_1 = [\mathbf{0}, \mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2], \qquad \tau_2 = -[\mathbf{0}, \mathbf{e}_2, \mathbf{e}_2 + \mathbf{e}_1].$$

Explicar por qué es razonable llamar a J^2 cuadrado unitario orientado positivamente en R^2. Mostrar que ∂J^2 es la suma de 4 1-simplex afines orientados. Determinarlos. ¿Qué es $\partial(\tau_1 - \tau_2)$?

18. Considérese el 3-simplex afín orientado

$$\sigma_1 = [\mathbf{0}, \mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3]$$

en R^3. Mostrar que σ_1 (considerado como una transformación lineal) tiene determinante 1. Entonces σ_1 está orientado positivamente.

Sean $\sigma_2, \ldots, \sigma_6$ otros cinco 3-simplex orientados, obtenidos de la siguiente manera: Hay cinco permutaciones (i_1, i_2, i_3) de $(1,2,3)$, distintas de $(1,2,3)$. Con cada (i_1, i_2, i_3) se asocia el simplex

$$s(i_1, i_2, i_3)[\mathbf{0}, \mathbf{e}_{i_1}, \mathbf{e}_{i_1} + \mathbf{e}_{i_2}, \mathbf{e}_{i_1} + \mathbf{e}_{i_2} + \mathbf{e}_{i_3}]$$

en donde s es el signo que aparece en la definición del determinante. (Asi es como τ_2 se obtuvo de τ_1 en el Ejercicio 17.)

Mostrar que $\sigma_2, \ldots, \sigma_6$ están orientados positivamente.

Si se pone $J^3 = \sigma_1 + \cdots + \sigma_6$, entonces J^3 puede denominarse cubo unitario orientado positivamente en R^3.

Mostrar que ∂J^3 es la suma de doce 2-simplex afines orientados. (Estos doce triángulos cubren la superficie del cubo unitario I^3.)

Mostrar que $\mathbf{x} = (x_1, x_2, x_3)$, está en el rango de σ_1 si, y solo si $0 \leq x_3, \leq x_2 \leq x_1 \leq 1$.

Mostrar que los rangos de $\sigma_2, \ldots, \sigma_6$ tienen interiores ajenos, y que su unión cubre I^3. (Compararlo con el Ejercicio 13; nótese que $3! = 6$.)

19. Sean J^2 y J^3 como los de los Ejercicios 17 y 18, respectivamente. Si se definen

$$\begin{aligned} B_{01}(u, v) &= (0, u, v), & B_{11}(u, v) &= (1, u, v), \\ B_{02}(u, v) &= (u, 0, v), & B_{12}(u, v) &= (u, 1, v), \\ B_{03}(u, v) &= (u, v, 0), & B_{13}(u, v) &= (u, v, 1). \end{aligned}$$

Estos son afines, y mapean R^2 en R^3.

Si se pone $\beta_{ri} = B_{ri}(J^2)$, para $r = 0, 1$, $i = 1,2,3$. Cada β_{ri} es una 2-cadena afín orientada. (Véase la Sec. 10.30.) De acuerdo al Ejercicio 18, verificar que

$$\partial J^3 = \sum_{i=1}^{3} (-1)^i (\beta_{0i} - \beta_{1i}),$$

20. Establecer las condiciones bajo las cuales, la fórmula

$$\int_{\Phi} f\, d\omega = \int_{\partial\Phi} f\omega - \int_{\Phi} (df) \wedge \omega$$

es válida, y mostrar que generaliza la fórmula de integración por partes.

Sugerencia: $d(f\omega) = (df) \wedge \omega + f\, d\omega$.

21. Como en el Ejercicio 10.36, considerar la 1-forma en $R^2 - \{\mathbf{0}\}$.

$$\eta = \frac{x\, dy - y\, dx}{x^2 + y^2}$$

(*a*) Efectuar el cálculo que conduce a la fórmula (113), y demostrar que $d\eta = 0$.

(*b*) Sea $\gamma(t) = (r \cos t, r \operatorname{sen} t)$, para algún $r > 0$, y Γ una curva-$\mathscr{C}''$ en $R^2 - \{\mathbf{0}\}$, con intervalo de parámetros $[0, 2\pi]$, y donde $\Gamma(0) = \Gamma(2\pi)$, tal que los intervalos $[\gamma(t), \Gamma(t)]$ no contienen a $\mathbf{0}$ para cualquier $t \in [0, 2\pi]$. Demostrar que

$$\int_{\Gamma} \eta = 2\pi.$$

Sugerencia: Para $0 \leq t \leq 2\pi$, $0 \leq u \leq 1$, defínase

$$\Phi(t, u) = (1 - u)\, \Gamma(t) + u\gamma(t).$$

Entonces Φ es una 2-superficie en $R^2 - \{\mathbf{0}\}$ cuyo dominio de parámetros es el rectángulo indicado. Debido a las cancelaciones (como en el Ejemplo 10.32),

$$\partial\Phi = \Gamma - \gamma.$$

Usar el teorema de Stokes para deducir que, debido a que $d\eta = 0$,

$$\int_{\Gamma} \eta = \int_{\gamma} \eta$$

(*c*) Si se toma $\Gamma(t) = (a \cos t, b \operatorname{sen} t)$ donde $a > 0$, $b > 0$ son fijos. Usar la parte (*b*) para mostrar que

$$\int_0^{2\pi} \frac{ab}{a^2 \cos^2 t + b^2 \operatorname{sen}^2 t}\, dt = 2\pi.$$

(*d*) Mostrar que en cualquier conjunto abierto en el cual $x \neq 0$,

$$\eta = d\left(\text{arc tan}\frac{y}{x}\right)$$

y que en cualquier conjunto abierto convexo en el cual $y \neq 0$,

$$\eta = d\left(-\text{arc tan}\frac{x}{y}\right)$$

Explicar por qué esto justifica la notación $\eta = d\theta$, a pesar de que η no es exacta en $R^2 - \{\mathbf{0}\}$.

(*e*) Mostrar que (*b*) puede deducirse de (*d*).

(*f*) Si Γ es cualquier curva-$\mathscr{C}'$ cerrada en $R^2 - \{\mathbf{0}\}$, demostrar que

$$\frac{1}{2\pi}\int_{\Gamma}\eta = \text{Ind}(\Gamma).$$

(Para la definición del índice de una curva, véase el Ejercicio 23 del Cap. 8.)

22. Se define ζ en $R^3 - \{\mathbf{0}\}$ como en el Ejemplo 10.37, por

$$\zeta = \frac{x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy}{r^3}$$

en donde $r = (x^2 + y^2 + z^2)^{1/2}$, y sea D el rectángulo dado por $0 \leq u \leq \pi$, $0 \leq v \leq 2\pi$, y Σ la 2-superficie en R^3, con dominio de parámetros D dado por

$$x = \text{sen}\, u \cos v, \qquad y = \text{sen}\, u\, \text{sen}\, v, \qquad z = \cos u.$$

(*a*) Demostrar que $d\zeta$ en $R^3 - \{\mathbf{0}\}$.

(*b*) Si S representa la restricción de Σ al dominio de parámetros $E \subset D$, demostrar que

$$\int_S \zeta = \int_E \text{sen}\, u\, du\, dv = A(S),$$

en donde A representa al área como en la sección 10.43. Nótese que ésta contiene como caso especial a (115).

(*c*) Supóngase que g, h_1, h_2, h_3, son funciones-$\mathscr{C}''$ sobre $[0, 1]$, $g > 0$. Sea $(x, y, z) = \Phi(s, t)$ una 2-superficie Φ, con dominio de parámetros I^2, definida por

$$x = g(t)h_1(s), \qquad y = g(t)h_2(s), \qquad z = g(t)h_3(s).$$

Demostrar directamente de (35), que

$$\int_{\Phi} \zeta = 0,$$

Nótese la forma del rango de Φ: Para s fijo, $\Phi(s, t)$ varía sobre un intervalo

en la línea que pasa por $\mathbf{0}$. Entonces el rango de Φ se encuentra en un "cono" con vértice en el origen.

(*d*) Sea E un rectángulo cerrado en D, con lados paralelos a los de D. Supóngase que $f \in \mathscr{C}''(D)$, $f > 0$. Sea Ω la 2-superficie con dominio de parámetros E definida por

$$\Omega(u, v) = f(u, v)\,\Sigma\,(u, v).$$

Definir S como en (*b*) y demostrar que

$$\int_{\Omega} \zeta = \int_{S} \zeta = A(S).$$

(Ya que S es la "proyección radial" de Ω en la esfera unitaria, este resultado hace que sea razonable llamar a $\int_{\Omega}\zeta$ el "ángulo sólido" subtendido por el rango de Ω en el origen.)

Sugerencia: Considerar la 3-superficie Ψ dada por

$$\Psi(t, u, v) = [1 - t + tf(u, v)]\,\Sigma\,(u, v),$$

en donde $(u, v) \in E$, $0 \leq t \leq 1$. Si v es fijo, el mapeo $(t, u) \rightarrow \Psi(t, u, v)$ es una 2-superficie Φ a la cual puede aplicarse (*c*) para demostrar que $\int_{\Phi}\zeta = 0$. Lo mismo se cumple cuando u es fijo. De (*a*) y el teorema de Stokes,

$$\int_{\partial\Psi} \zeta = \int_{\Psi} d\zeta = 0.$$

(*e*) Hacer $\lambda = -(z/r)\eta$, en donde como en el Ejercicio 21,

$$\eta = \frac{x\,dy - y\,dx}{x^2 + y^2},$$

Entonces γ es una 1-forma en el conjunto abierto $V \subset R^3$ en el cual $x^2 + y^2 > 0$. Mostrar que ζ es *exacta en* V probando que

$$\zeta = d\lambda.$$

(*f*) A partir de (*e*) deducir (*d*), sin usar (*c*).

Sugerencia: Para empezar, suponer que $0 < u < \pi$ sobre E. Por (*e*) se tiene,

$$\int_{\Omega} \zeta = \int_{\partial\Omega} \lambda \qquad \text{y} \qquad \int_{S} \zeta = \int_{\partial S} \lambda.$$

Mostrar que las dos integrales de λ son iguales, usando la parte (*d*) del Ejercicio 21 y notando que z/r es lo mismo en $\Sigma(u, v)$ que en $\Omega(u, v)$.

(*g*) ¿Es exacta ζ en el complemento de cada recta que pasa por el origen?

23. Si n es fijo, y se define $r_k = (x_1^2 + \cdots + x_k^2)^{1/2}$ para $1 \leq k \leq n$, sea E_k el con-

junto de todos los $\mathbf{x} \in R^n$ en los cuales $r_k > 0$, y ω_k la $(k-1)$-forma definida en E_k por

$$\omega_k = (r_k)^{-k} \sum_{i=1}^{k} (-1)^{i-1} x_i \, dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_k.$$

Nótese que con la terminología de los Ejercicios 21 y 22, $\omega_2 = \eta$, $\omega_3 = \zeta$. Nótese también que

$$E_1 \subset E_2 \subset \cdots \subset E_n = R^n - \{\mathbf{0}\}.$$

(*a*) Demostrar que $d\omega_k = 0$ en E_k.

(*b*) Para $k = 2, \ldots, n$, demostrar que ω_k es exacta en E_{k-1}, mostrando que

$$\omega_k = d(f_k \omega_{k-1}) = (df_k) \wedge \omega_{k-1},$$

en donde $f_k(\mathbf{x}) = (-1)^k g_k(x_k/r_k)$, y

$$g_k(t) = \int_{-1}^{t} (1 - s^2)^{(k-3)/2} \, ds \qquad (-1 < t < 1).$$

Sugerencia: f_k satisface las ecuaciones diferenciales

$$\mathbf{x} \cdot (\nabla f_k)(\mathbf{x}) = 0$$

y

$$(D_k f_k)(\mathbf{x}) = \frac{(-1)^k (r_{k-1})^{k-1}}{(r_k)^k}.$$

(*c*) ¿Es exacta ω_n en E_n?

(*d*) Nótese que (*b*) es una generalización de la parte (*e*) del Ejercicio 22. Intentar ampliar alguna de las otras afirmaciones de los Ejercicios 21 y 22 para ω_n, con n arbitrario.

24. Sea $\omega = \Sigma a_i(\mathbf{x}) \, dx_i$ una 1-forma de clase $\mathscr{C}''$ en un conjunto abierto convexo $E \subset R^n$. Admitir que $d\omega = 0$, y demostrar que ω es exacta en E, completando el siguiente plan:

Se fija $\mathbf{p} \in E$, y se define

$$f(\mathbf{x}) = \int_{[\mathbf{p},\mathbf{x}]} \omega \qquad (\mathbf{x} \in E).$$

Aplicar el teorema de Stokes a los 2-simplex afines orientados $[\mathbf{p},\mathbf{x},\mathbf{y}]$ en E. Deducir que

$$f(\mathbf{y}) - f(\mathbf{x}) = \sum_{i=1}^{n} (y_i - x_i) \int_0^1 a_i((1-t)\mathbf{x} + t\mathbf{y}) \, dt$$

para $\mathbf{x} \in E$, $\mathbf{y} \in E$. De lo anterior concluir que $(D_i f)(\mathbf{x}) = a_i(\mathbf{x})$.

25. Suponer que ω es una 1-forma en un conjunto abierto $E \subset R^n$ tal que

$$\int_\gamma \omega = 0$$

para cada curva cerrada γ en E, de clase $\mathscr{C}'$. Demostrar que ω es exacta en E, imitando parte del argumento planteado en el Ejercicio 24.

26. Suponer que ω es una 1-forma en $R^3 - \{\mathbf{0}\}$, de clase $\mathscr{C}'$ y $d\omega = 0$. Demostrar que ω es exacta en $R^3 - \{\mathbf{0}\}$.

Sugerencia: Cada curva cerrada continuamente diferenciable en $R^3 - \{\mathbf{0}\}$ es la frontera de una 2-superficie en $R^3 - \{\mathbf{0}\}$. Aplicar el teorema de Stokes y el Ejercicio 25.

27. Sea E una 3-celda abierta en R^3, con lados paralelos a los ejes coordenados. Suponer que $(a,b,c) \in E, f_i \in \mathscr{C}'(E)$ para $i = 1,2,3$,

$$\omega = f_1\, dy \wedge dz + f_2\, dz \wedge dx + f_3\, dx \wedge dy,$$

y admitir que $d\omega = 0$ en E. Si se define

$$\lambda = g_1\, dx + g_2\, dy$$

en donde

$$g_1(x, y, z) = \int_c^z f_2(x, y, s)\, ds - \int_b^y f_3(x, t, c)\, dt$$

$$g_2(x, y, z) = -\int_c^z f_1(x, y, s)\, ds,$$

para $(x,y,z) \in E$. Demostrar que $d\lambda = \omega$ en E.

Evaluar estas integrales cuando $\omega = \zeta$ y encontrar después la forma λ que ocurre en la parte (e) del Ejercicio 22.

28. Con $b > a > 0$ fijos, se define

$$\Phi(r, \theta) = (r \cos \theta, r \operatorname{sen} \theta)$$

para $a \leq r \leq b$, $0 \leq \theta \leq 2\pi$. (El rango de Φ es un anillo en R^2.) Hacer $\omega = x^3\, dy$ y calcular

$$\int_\Phi d\omega \qquad \text{y} \qquad \int_{\partial\Phi} \omega$$

para verificar que son iguales.

29. Demostrar la existencia de una función α con las propiedades necesarias en la demostración del Teorema 10.38, y demostrar también que la función resultante F es de clase $\mathscr{C}'$. (Ambas afirmaciones se vuelven triviales si E es una celda abierta o una bola abierta, debido a que α entonces puede tomarse como una constante. Referirse al Teorema 9.42.)

30. Si $\mathbf{N}$ es el vector dado en (135), demostrar que

$$det \begin{bmatrix} \alpha_1 & \beta_1 & \alpha_2\beta_3 - \alpha_3\beta_2 \\ \alpha_2 & \beta_2 & \alpha_3\beta_1 - \alpha_1\beta_3 \\ \alpha_3 & \beta_3 & \alpha_1\beta_2 - \alpha_2\beta_1 \end{bmatrix} = |\mathbf{N}|^2.$$

Verificar también la ecuación (137).

31. Sea $E \subset R^3$ abierto, supóngase $g \in \mathscr{C}''(E)$, $h \in \mathscr{C}''(E)$ y considérese el campo vectorial

$$\mathbf{F} = g \nabla h.$$

(*a*) Demostrar que

$$\nabla \cdot F = g \nabla^2 h + (\nabla g) \cdot (\nabla h)$$

en donde $\nabla^2 h = \nabla \cdot (\nabla h) = \Sigma \partial^2 h/\partial x_i^2$ se llama el "Laplaciano" de h.

(*b*) Si Ω es un subconjunto cerrado de E con frontera orientada positivamente (como en el Teorema 10.51), demostrar que

$$\int_\Omega [g \nabla^2 h + (\nabla g) \cdot (\nabla h)] dV = \int_{\partial\Omega} g \frac{\partial h}{\partial n} dA$$

en donde (como de costumbre) se escribió $\partial h/\partial n$ en lugar de $(\nabla h) \cdot \mathbf{n}$. (Por lo anterior, $\partial h/\partial n$ es la derivada direccional de h en la dirección de la normal hacia afuera a $\partial\Omega$, denominada *derivada normal* de h.) Intercambiar g y h, restar la fórmula resultante de la primera, para obtener

$$\int_\Omega (g \nabla^2 h - h \nabla^2 g)\, dV = \int_{\partial\Omega} \left(g \frac{\partial h}{\partial n} - h \frac{dg}{\partial n} \right) dA.$$

Estas dos fórmulas se llaman comúnmente *identidades de Green.*

(*c*) Admitir que h es *armónica* en E ; esto significa que $\nabla^2 h = 0$. Tomar $g = 1$ y concluir que

$$\int_{\partial\Omega} \frac{\partial h}{\partial n} dA = 0.$$

Tomar $g = h$ y concluir que $h = 0$ en Ω si $h = 0$ sobre $\partial\Omega$.

(*d*) Mostrar que las identidades de Green son válidas también en R^2.

32. Con δ, $0 < \delta < 1$ fijo, sea D el conjunto de todos los $(\theta, t) \in R^2$ tales que $0 \leq \theta \leq \pi$, $-\delta \leq t \leq \delta$. Sea Φ la 2-superficie en R^3, con dominio de parámetros D, dada por

$$x = (1 - t \operatorname{sen} \theta) \cos 2\theta$$
$$y = (1 - t \operatorname{sen} \theta) \operatorname{sen} 2\theta$$
$$z = t \cos \theta$$

en donde $(x,y,z) = \Phi(\theta, t)$. Nótese que $\Phi(\pi, t) = \Phi(0, -t)$ y que Φ es uno-a-uno en el resto de D.

El rango $M = \Phi(D)$ de Φ se conoce con el nombre de *Cinta de Möbius*. Este es el ejemplo más sencillo de una superficie no orientable.

Demostrar las distintas afirmaciones que se hacen en la descripción siguiente: Poner $\mathbf{p}_1 = (0, -\delta)$, $\mathbf{p}_2 = (\pi, -\delta)$, $p_3 = (\pi, \delta)$ $\mathbf{p}_4 = (0, \delta)$, $\mathbf{p}_5 = \mathbf{p}_1$. Poniendo $\gamma_i = [\mathbf{p}_i, \mathbf{p}_{i+1}]$, $i = 1, \ldots, 4$, y $\Gamma_i = \Phi \circ \gamma_i$. Entonces

$$\partial\Phi = \Gamma_1 + \Gamma_2 + \Gamma_3 + \Gamma_4.$$

Poniendo $\mathbf{a} = (1, 0, -\delta)$, $\mathbf{b} = (1, 0, \delta)$. Entonces

$$\Phi(\mathbf{p}_1) = \Phi(\mathbf{p}_3) = \mathbf{a}, \qquad \Phi(\mathbf{p}_2) = \Phi(\mathbf{p}_4) = \mathbf{b},$$

y $\partial\Phi$ se puede describir como sigue.

Γ_1 avanza en espiral hacia arriba de $\mathbf{a}$ a $\mathbf{b}$; su proyección en el plano (x,y) tiene $+1$ como número de arrollamiento alrededor del origen. (Véase Ejercicio 23 del Cap. 8.)

$\Gamma_2 = [\mathbf{b}, \mathbf{a}]$.

Γ_3 avanza en espiral hacia arriba de $\mathbf{a}$ a $\mathbf{b}$; su proyección en el plano (x,y) tiene -1 como número de arrollamiento alrededor del origen.

$\Gamma_4 = [\mathbf{b}, \mathbf{a}]$.

Por lo anterior, $\partial\Phi = \Gamma_1 + \Gamma_3 + 2\Gamma_2$.

Si vamos de $\mathbf{a}$ a $\mathbf{b}$ a lo largo de Γ_1 y continuamos a lo largo de la "cara" de M hasta que regresamos a $\mathbf{a}$, la curva trazada es

$$\Gamma = \Gamma_1 - \Gamma_3,$$

la cual también puede representarse sobre el intervalo de parámetros $[0, 2\pi]$ por medio de las ecuaciones

$$\begin{aligned} x &= (1 + \delta \operatorname{sen} \theta) \cos 2\theta \\ y &= (1 + \delta \operatorname{sen} \theta) \operatorname{sen} 2\theta \\ z &= -\delta \cos \theta. \end{aligned}$$

Debe subrayarse que $\Gamma \neq \partial\Phi$: Sea η la 1-forma que se vio en los Ejercicios 21 y 22. Como $d\eta = 0$, el teorema de Stokes muestra que

$$\int_{\partial\Phi} \eta = 0.$$

Pero aunque Γ es la frontera "geométrica" de M, se tiene que

$$\int_{\Gamma} \eta = 4\pi.$$

Con el fin de evitar esta posible fuente de confusión, con frecuencia se establece la fórmula de Stokes (Teorema 10.50) sólo para superficies orientables Φ.

11

TEORÍA DE LEBESGUE

El propósito de este capítulo es presentar los conceptos fundamentales de la teoría de la medida y la integración de Lebesgue y demostrar algunos de sus teoremas principales en una forma de exposición bastante general, para no enmascarar las líneas fundamentales del desarrollo en un conjunto de detalles de menor importancia. Por ello, en varios casos las demostraciones están solo esbozadas, y algunas de las proposiciones más sencillas se enuncian sin demostración. Sin embargo, el lector que se ha familiarizado con las técnicas utilizadas en los capítulos precedentes, no hallará ninguna dificultad para llenar los pasos que faltan.

La teoría de la integral de Lebesgue puede ser desarrollada de varios modos distintos. Solo estudiaremos aquí uno de esos métodos. Para otros procedimientos, consúltense los tratados más especializados en integración, reseñados en la Bibliografía.

FUNCIONES DE CONJUNTOS

Si A y B son dos conjuntos cualesquiera, escribimos $A - B$ para el conjunto de todos los elementos x, tales que $x \in A$, $x \notin B$. La notación $A - B$ no implica que $B \subset A$. Expresamos el conjunto vacío por 0. y decimos que A y B son ajenos si $A \cap B = 0$.

11.1 Definición Se dice que una familia $\mathscr{R}$ de conjuntos es un *anillo* si $A \in \mathscr{R}$ y $B \in \mathscr{R}$ implica

(1) $$A \cup B \in \mathscr{R}, \qquad A - B \in \mathscr{R}.$$

Como $A \cap B = A - (A - B)$, tenemos también que $A \cap B \in \mathscr{R}$ si $\mathscr{R}$ es un anillo.

A un anillo $\mathscr{R}$ se le llama *σ-anillo* si

(2) $$\bigcup_{n=1}^{\infty} A_n \in \mathscr{R}$$

siempre que $A_n \in \mathscr{R}$ $(n = 1, 2, 3, \ldots)$. Como

$$\bigcap_{n=1}^{\infty} A_n = A_1 - \bigcup_{n=1}^{\infty} (A_1 - A_n),$$

tenemos también que

$$\bigcap_{n=1}^{\infty} A_n \in \mathscr{R}$$

si $\mathscr{R}$ es un σ-anillo.

11.2 Definición Diremos que ϕ es una función de conjuntos definida en $\mathscr{R}$ si ϕ asigna a cada $A \in \mathscr{R}$ un número $\phi(A)$ del sistema ampliado de los números reales. ϕ es *aditiva* si $A \cap B = 0$ implica

(3) $$\phi(A \cup B) = \phi(A) + \phi(B),$$

y es *aditiva numerable* si $A_i \cap A_j = 0$ $(i \neq j)$ implica

(4) $$\phi\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \phi(A_n).$$

Supondremos siempre que el rango de ϕ no contiene a $+\infty$ y $-\infty$; porque si los contuviera, el segundo miembro de (3) carecería de sentido. También excluimos las funciones de conjuntos, cuyos únicos valores son $+\infty$ o $-\infty$.

Es interesante observar que el primer miembro de (4) es independiente del orden en que estén colocadas las A_n. Por tanto, el teorema de los reordenamientos demuestra que el segundo miembro de (4) converge absolutamente si converge de algún modo; si no converge, las sumas parciales tienden a $+\infty$ o a $-\infty$.

Si ϕ es aditiva, se comprueban fácilmente las siguientes propiedades

(5) $$\phi(0) = 0.$$

(6) $$\phi(A_1 \cup \cdots \cup A_n) = \phi(A_1) + \cdots + \phi(A_n)$$

si $A_i \cap A_j = 0$ siempre que $i \neq j$.

(7) $$\phi(A_1 \cup A_2) + \phi(A_1 \cap A_2) = \phi(A_1) + \phi(A_2).$$

Si $\phi(A) \geq 0$ para todo A, y $A_1 \subset A_2$, será

(8) $$\phi(A_1) \leq \phi(A_2).$$

Como consecuencia de (8), las funciones de conjuntos aditivas no negativas son llamadas a veces monótonas.

(9) $$\phi(A - B) = \phi(A) - \phi(B)$$

si $B \subset A$, y $|\phi(B)| < +\infty$.

11.3 Teorema *Supongamos que ϕ es aditiva numerable en un anillo $\mathscr{R}$, que $A_n \in \mathscr{R}$ $(n = 1,2,3,\ldots)$, $A_1 \subset A_2 \subset A_3 \subset \cdots$, $A \in \mathscr{R}$, y*

$$A = \bigcup_{n=1}^{\infty} A_n.$$

Entonces, cuando $n \to \infty$.

$$\phi(A_n) \to \phi(A).$$

Demostración Pongamos $B_1 = A_1$, y

$$B_n = A_n - A_{n-1} \qquad (n = 2, 3, \ldots).$$

Entonces $B_i \cap B_j = 0$ para $i \neq j$, $A_n = B_1 \cup \cdots \cup B_n$ y $A = \bigcup B_n$. Por tanto,

$$\phi(A_n) = \sum_{i=1}^{n} \phi(B_i)$$

y

$$\phi(A) = \sum_{i=1}^{\infty} \phi(B_i).$$

CONSTRUCCIÓN DE LA MEDIDA DE LEBESGUE

11.4 Definición Representemos por R^p un espacio euclidiano p-dimensional. Por *intervalo* en R^p entenderemos el conjunto de los puntos $\mathbf{x} = (x_1, \ldots, x_p)$, tales que

(10) $$a_i \leq x_i \leq b_i \qquad (i = 1, \ldots, p),$$

o el conjunto de puntos caracterizado por (10) con alguno o todos los signos $\leq$ sustituido por $<$. La posibilidad de ser $a_i = b_i$ para algún valor de i no queda excluida; en particular, el conjunto vacío está incluido entre los intervalos.

Si A es la unión de un número finito de intervalos, se dice que A es un *conjunto elemental.*

Si I es un intervalo, definimos

$$m(I) = \prod_{i=1}^{p} (b_i - a_i),$$

sin que importe si en cualquiera de las desigualdades de (10) está excluida o incluida la igualdad.

Si $A = I_1 \cup \cdots \cup I_n$, y si estos intervalos son ajenos dos a dos, pondremos

(11) $$m(A) = m(I_1) + \cdots + m(I_n).$$

Representaremos por $\mathscr{E}$ la familia de todos los subconjuntos elementales de R^p.

Podrían comprobarse ahora las siguientes propiedades:

(12) $\mathscr{E}$ es un anillo, pero no un σ-anillo.
(13) Si $A \in \mathscr{E}$, A es la unión de un número finito de intervalos *ajenos*.
(14) Si $A \in \mathscr{E}$, $m(A)$ está bien definida por (11); esto es, si se utilizan dos descomposiciones diferentes de A en intervalos ajenos, las dos dan lugar al mismo valor de $m(A)$.
(15) m es aditiva en $\mathscr{E}$.

Nótese que si $p = 1,2,3$, m es longitud, área y volumen, respectivamente.

11.5 Definición Se dice que una función de conjuntos aditiva no negativa ϕ, definida en $\mathscr{E}$, es *regular* si es cierto lo siguiente: Para cada $A \in \mathscr{E}$ y cada $\varepsilon > 0$, existen conjuntos $F \in \mathscr{E}$, $G \in \mathscr{E}$, tales que F es cerrado, G es abierto, $F \subset A \subset G$ y

(16) $$\phi(G) - \varepsilon \leq \phi(A) \leq \phi(F) + \varepsilon.$$

11.6 Ejemplos

(*a*) *La función de conjuntos m es regular.*

Si A es un intervalo, es evidente que se satisfacen las condiciones de la Definición 11.5. El caso general se deduce de (13).

(*b*) Tomemos $R^p = R^1$, y sea α una función monótona creciente, definida para todo x real. Pongamos

$$\mu([a, b)) = \alpha(b-) - \alpha(a-),$$

$$\mu([a, b]) = \alpha(b+) - \alpha(a-),$$

$$\mu((a, b]) = \alpha(b+) - \alpha(a+),$$

$$\mu((a, b)) = \alpha(b-) - \alpha(a+).$$

Donde $[a,b)$ es el conjunto $a \leq x < b$, etc. Hay que distinguir estos casos a causa de las posibles discontinuidades de α. Si μ está definido para conjuntos elementales del modo hecho en (11), es regular en $\mathscr{E}$. La demostración es igual que la anterior de (*a*).

Nuestro próximo objetivo es demostrar que toda función de conjuntos regular en $\mathscr{E}$ puede ser extendida a una función de conjuntos aditiva numerable en un σ-anillo que contiene a $\mathscr{E}$.

11.7 Definición Sea μ aditivo, regular, no negativo y finito en $\mathscr{E}$. Consideremos cubiertas numerables de cualquier conjunto $E \subset R^p$ por conjuntos abiertos elementales A_n:

$$E \subset \bigcup_{n=1}^{\infty} A_n .$$

Definamos

$$\mu^*(E) = \inf \sum_{n=1}^{\infty} \mu(A_n), \tag{17}$$

tomando el inf sobre todas las cubiertas numerables de E por conjuntos elementales abiertos. A $\mu^*(E)$ se le llama *medida exterior* de E correspondiente a μ.

Es evidente que $\mu^*(E) \geq 0$ para todo E y que

$$\mu^*(E_1) \leq \mu^*(E_2) \tag{18}$$

si $E_1 \subset E_2$.

11.8 Teorema

(*a*) *Para todo* $A \in \mathscr{E}$, $\mu^*(A) = \mu(A)$.

(*b*) *Si* $E = \bigcup_1^\infty E_n$,

(19) $$\mu^*(E) \le \sum_{n=1}^{\infty} \mu^*(E_n).$$

Nótese que (*a*) afirma que μ^* es una prolongación de μ, de $\mathscr{E}$ a la familia de *todos* los subconjuntos de R^p. A la propiedad (19) se le llama *subaditividad.*

Demostración Elijamos $A \in \mathscr{E}$ y $\varepsilon > 0$.

La regularidad de μ^* demuestra que A está contenido en un conjunto elemental abierto G tal que $\mu(G) \le \mu(A) + \varepsilon$. Como $\mu^*(A) \le \mu(G)$ y como ε era arbitrario, tenemos

(20) $$\mu^*(A) \le \mu(A).$$

La definición de μ^* demuestra que hay una sucesión $\{A_n\}$ de conjuntos elementales abiertos cuya unión contiene a A, tal que

$$\sum_{n=1}^{\infty} \mu(A_n) \le \mu^*(A) + \varepsilon.$$

La regularidad de μ demuestra que A contiene un conjunto elemental cerrado F tal que $\mu(F) \ge \mu(A) - \varepsilon$; y como F es compacto, tenemos

$$F \subset A_1 \cup \cdots \cup A_N$$

para algún N. Por tanto,

$$\mu(A) \le \mu(F) + \varepsilon \le \mu(A_1 \cup \cdots \cup A_N) + \varepsilon \le \sum_1^N \mu(A_n) + \varepsilon \le \mu^*(A) + 2\varepsilon.$$

Lo que, en unión de (20), demuestra (*a*).

Continuando, supongamos que $E = \bigcup E_n$ y admitamos que $\mu^*(E_n) < +\infty$ para todo n. Dado $\varepsilon > 0$, hay cubiertas $\{A_{nk}\}$, $k = 1,2,3,\ldots$, de E_n por conjuntos elementales abiertos tales que

(21) $$\sum_{k=1}^{\infty} \mu(A_{nk}) \le \mu^*(E_n) + 2^{-n}\varepsilon.$$

Entonces

$$\mu^*(E) \le \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu(A_{nk}) \le \sum_{n=1}^{\infty} \mu^*(E_n) + \varepsilon,$$

y se deduce (19). En el caso excluido, esto es, si $\mu^*(E_n) = +\infty$ para algún n, (19) es trivial.

11.9 Definición Para cada $A \subset R^p$, $B \subset R^p$, definiremos

(22) $$S(A, B) = (A - B) \cup (B - A),$$

(23) $$d(A, B) = \mu^*(S(A, B)).$$

Escribiremos $A_n \to A$ si

$$\lim_{n\to\infty} d(A, A_n) = 0.$$

Si hay una sucesión $\{A_n\}$ de conjuntos elementales tales que $A_n \to A$, diremos que A es *μ-medible finitamente* y escribiremos $A \in \mathfrak{M}_F(\mu)$.

Si A es la unión de una colección numerable de conjuntos μ-medibles finitamente, diremos que A es *μ-medible* y escribiremos $A \in \mathfrak{M}(\mu)$.

$S(A,B)$ es la llamada «diferencia simétrica» de A y B. Veremos que $d(A,B)$ es esencialmente una función distancia.

El teorema siguiente nos permitirá obtener la prolongación deseada de μ.

11.10 Teorema *$\mathfrak{M}(\mu)$ es un σ-anillo y μ^* es aditiva numerable en $\mathfrak{M}(\mu)$.*

Antes de volver a la demostración de este teorema, desarrollaremos algunas de las propiedades de $S(A,B)$ y $d(A,B)$. Tenemos

(24) $$S(A, B) = S(B, A), \qquad S(A, A) = 0.$$

(25) $$S(A, B) \subset S(A, C) \cup S(C, B).$$

(26) $$\left.\begin{array}{r} S(A_1 \cup A_2, B_1 \cup B_2) \\ S(A_1 \cap A_2, B_1 \cap B_2) \\ S(A_1 - A_2, B_1 - B_2) \end{array}\right\} \subset S(A_1, B_1) \cup S(A_2, B_2).$$

(24) es clara, y (25) se deduce de

$$(A - B) \subset (A - C) \cup (C - B), \qquad (B - A) \subset (C - A) \cup (B - C).$$

La primera fórmula de (26) se obtiene de

$$(A_1 \cup A_2) - (B_1 \cup B_2) \subset (A_1 - B_1) \cup (A_2 - B_2).$$

Ahora, escribiendo E^c para el complemento de E, tenemos

$$S(A_1 \cap A_2, B_1 \cap B_2) = S(A_1^c \cup A_2^c, B_1^c \cup B_2^c)$$
$$\subset S(A_1^c, B_1^c) \cup S(A_2^c, B_2^c) = S(A_1, B_1) \cup S(A_2, B_2);$$

y se obtiene la última fórmula de (26) si observamos que

$$A_1 - A_2 = A_1 \cap A_2^c.$$

Por (23), (19) y (18), estas propiedades de $S(A,B)$ implican

(27) $$d(A, B) = d(B, A), \qquad d(A, A) = 0,$$

(28) $$d(A, B) \leq d(A, C) + d(C, B),$$

(29) $$\left.\begin{array}{l} d(A_1 \cup A_2, B_1 \cup B_2) \\ d(A_1 \cap A_2, B_1 \cap B_2) \\ d(A_1 - A_2, B_1 - B_2) \end{array}\right\} \leq d(A_1, B_1) + d(A_2, B_2).$$

Las relaciones (27) y (28) muestran que $d(A,B)$ satisface las condiciones de la Definición 2.15, excepto que $d(A,B) = 0$ no implica $A = B$. Por ejemplo, si $\mu = m$, A es numerable, y B es vacío, tenemos

$$d(A, B) = m^*(A) = 0;$$

para verlo, cubrimos el n-ésimo punto de A por un intervalo I_n tal que

$$m(I_n) < 2^{-n}\varepsilon.$$

Pero si definimos dos conjuntos A y B como equivalentes cuando

$$d(A, B) = 0,$$

dividimos los subconjuntos de R^p en clases de equivalencia, y $d(A,B)$ forma el conjunto de estas clases de equivalencia en un espacio métrico. $\mathfrak{M}_F(\mu)$ se obtiene entonces, como cerradura de $\mathscr{E}$. Esta interpretación no es esencial para la demostración, pero explica la idea que sirve de base.

Necesitamos otra propiedad más de $d(A,B)$, esto es,

(30) $$|\mu^*(A) - \mu^*(B)| \leq d(A, B),$$

si al menos una de $\mu^*(A)$, $\mu^*(B)$ es finita. Porque suponiendo que $0 \leq \mu^*(B) \leq \mu^*(A)$. Entonces (28) muestra que

$$d(A,0) \leq d(A, B) + d(B, 0),$$

esto es,

$$\mu^*(A) \le d(A, B) + \mu^*(B).$$

Como $\mu^*(B)$ es finita, se deduce que

$$\mu^*(A) - \mu^*(B) \le d(A, B).$$

Demostración del Teorema 11.10 Supongamos que $A \in \mathfrak{M}_F(\mu)$, $B \in \mathfrak{M}_F(\mu)$. Elijamos $\{A_n\}$, $\{B_n\}$ de modo que $A_n \in \mathscr{E}$. $B_n \in \mathscr{E}$, $A_n \to A$, $B_n \to B$. Entonces (29) y (30) demuestran que

(31) $$A_n \cup B_n \to A \cup B,$$

(32) $$A_n \cap B_n \to A \cap B,$$

(33) $$A_n - B_n \to A - B,$$

(34) $$\mu^*(A_n) \to \mu^*(A),$$

y $\mu^*(A) < +\infty$, pues $d(A_n, A) \to 0$. Por (31) y (33), $\mathfrak{M}_F(\mu)$ es un anillo. Por (7),

$$\mu(A_n) + \mu(B_n) = \mu(A_n \cup B_n) + \mu(A_n \cap B_n).$$

Suponiendo $n \to \infty$, obtenemos, por (34) y el Teorema 11.8(a),

$$\mu^*(A) + \mu^*(B) = \mu^*(A \cup B) + \mu^*(A \cap B).$$

Si $A \cap B = 0$, entonces $\mu^*(A \cap B) = 0$.

Se deduce que μ^* es aditiva en $\mathfrak{M}_F(\mu)$.

Sea, ahora, $A \in \mathfrak{M}(\mu)$. Entonces puede representarse A como la unión de una colección numerable de conjuntos *ajenos* de $\mathfrak{M}_F(\mu)$; porque si $A = \bigcup A_n'$, con $A_n' \in \mathfrak{M}_F(\mu)$, escribiremos $A_1 = A_1'$ y

$$A_n = (A_1' \cup \cdots \cup A_n') - (A_n' \cup \cdots \cup A_{n-1}') \qquad (n = 2, 3, 4, \ldots).$$

Entonces

(35) $$A = \bigcup_{n=1}^{\infty} A_n$$

es la representación apropiada. Por (19)

(36) $$\mu^*(A) \le \sum_{n=1}^{\infty} \mu^*(A_n).$$

Por otro lado, $A \supset A_1 \cup \cdots \cup A_n$, y por la aditividad de μ^* en $\mathfrak{M}_F(\mu)$ obtenemos

(37) $$\mu^*(A) \geq \mu^*(A_1 \cup \cdots \cup A_n) = \mu^*(A_1) + \cdots + \mu^*(A_n).$$

Las ecuaciones (36) y (37) implican

(38) $$\mu^*(A) = \sum_{n=1}^{\infty} \mu^*(A_n).$$

Supongamos que $\mu^*(A)$ es finita. Pongamos $B_n = A_1 \cup \cdots \cup A_n$. Entonces, (38) demuestra que

$$d(A, B_n) = \mu^*\left(\bigcup_{i=n+1}^{\infty} A_i\right) = \sum_{i=n+1}^{\infty} \mu^*(A_i) \to 0$$

cuando $n \to \infty$. Por tanto $B_n \to A$, y como $B_n \in \mathfrak{M}_F(\mu)$, se ve fácilmente que $A \in \mathfrak{M}_F(\mu)$.

Hemos demostrado así que $A \in \mathfrak{M}_F(\mu)$ si $A \in \mathfrak{M}(\mu)$ y $\mu^*(A) < +\infty$.

Está claro ahora que μ^* es aditiva en forma numerable en $\mathfrak{M}(\mu)$, porque si

$$A = \bigcup A_n,$$

donde $\{A_n\}$ es una sucesión de conjuntos ajenos de $\mathfrak{M}(\mu)$, hemos demostrado que (38) se cumple si $\mu^*(A_n) < +\infty$ para todo n, y en caso contrario (38) es evidente.

Finalmente, tenemos que demostrar que $\mathfrak{M}(\mu)$, es un σ-anillo. Si $A_n \in \mathfrak{M}(\mu)$, $n = 1,2,3,\ldots$, es evidente que $\bigcup A_n \in \mathfrak{M}(\mu)$ (Teorema 2.12). Supongamos que $A \in \mathfrak{M}(\mu)$, $B \in \mathfrak{M}(\mu)$, y

$$A = \bigcup_{n=1}^{\infty} A_n, \qquad B = \bigcup_{n=1}^{\infty} B_n,$$

donde $A_n, B_n \in \mathfrak{M}_F(\mu)$. Entonces, la identidad

$$A_n \cap B = \bigcup_{i=1}^{\infty} (A_n \cap B_i)$$

demuestra que $A_n \cap B \in \mathfrak{M}(\mu)$; y como

$$\mu^*(A_n \cap B) \leq \mu^*(A_n) < +\infty,$$

$A_n \cap B \in \mathfrak{M}_F(\mu)$. Por tanto $A_n - B \in \mathfrak{M}_F(\mu)$, y $A - B \in \mathfrak{M}(\mu)$, pues $A - B = \bigcup_{n=1}^{\infty} (A_n - B)$.

Sustituiremos ahora $\mu^*(A)$ por $\mu(A)$ si $A \in \mathfrak{M}(\mu)$. Así, μ definido originalmente solo en $\mathscr{E}$, queda prolongado a una función de conjuntos aditiva en forma numerable en el σ-anillo $\mathfrak{M}(\mu)$. Esta función de conjuntos prolongada se llama *medida*. El caso particulár $\mu = m$ es la llamada medida de Lebesgue en R^p.

11.11 Observaciones

(*a*) Si A es abierto, entonces $A \in \mathfrak{M}(\mu)$ porque cada conjunto abierto en R^p es la unión de una colección numerable de intervalos abiertos. Para verlo, es suficiente construir una base numerable cuyos elementos sean intervalos abiertos.

Tomando complementos, se deduce que todo conjunto cerrado está en $\mathfrak{M}(\mu)$.

(*b*) Si $A \in \mathfrak{M}(\mu)$ y $\varepsilon > 0$, existen conjuntos F y G tales que

$$F \subset A \subset G,$$

F es cerrado, G es abierto, y

(39) $$\mu(G - A) < \varepsilon, \qquad \mu(A - F) < \varepsilon.$$

La primera desigualdad se cumple porque μ^* fue definida por medio de recubrimientos por conjuntos elementales *abiertos*. La segunda se deduce tomando complementos.

(*c*) Decimos que E es un *conjunto de Borel*, si E puede obtenerse por número de operaciones numerable, partiendo de conjuntos abiertos, consistiendo cada operación en tomar uniones, intersecciones o complementos. La colección $\mathscr{B}$ de todos los conjuntos de Borel en R^p es un σ-anillo; de hecho, es el menor σ-anillo que contiene todos los conjuntos abiertos. Por la observación (*a*), $E \in \mathfrak{M}(\mu)$ si $E \in \mathscr{B}$.

(*d*) Si $A \in \mathfrak{M}(\mu)$, existen conjuntos de Borel F y G tales que $F \subset A \subset G$, y

(40) $$\mu(G - A) = \mu(A - F) = 0.$$

Lo que se deduce de (*b*) si tomamos $\varepsilon = 1/n$ y hacemos $n \to \infty$.

Como $A = F \cup (A - F)$, vemos que cada $A \in \mathfrak{M}(\mu)$ es la unión de un conjunto de Borel y un conjunto de medida cero.

Los conjuntos de Borel son μ-medibles para todo μ. Pero los conjuntos de medida cero [esto es, los conjuntos E para los cuales $\mu^*(E) = 0$] pueden ser diferentes para diferentes μ.

(*e*) Para cada μ, los conjuntos de medida cero forman un σ-anillo.

(*f*) En el caso de la medida de Lebesgue, todo conjunto numerable

tiene medida cero. Pero hay conjuntos no numerables (de hecho, perfectos) de medida cero. Puede tomarse como ejemplo el conjunto de Cantor: Utilizando la notación de la sección 2.44, se ve fácilmente que

$$m(E_n) = (\tfrac{2}{3})^n \qquad (n = 1, 2, 3, \ldots);$$

y como $P = \bigcap E_n$, $P \subset E_n$ para todo n, de modo que $m(P) = 0$.

ESPACIOS DE MEDIDA

11.12 Definición Supongamos que X es un conjunto, no necesariamente subconjunto de un espacio euclidiano, o en realidad de algún espacio métrico. Se dice que X es un *espacio de medida* si existen un σ-anillo de subconjuntos de X (llamados conjuntos medibles) y una función de conjuntos aditiva numerable no negativa μ (llamada medida), definida en $\mathfrak{M}$.

Si, además, $X \in \mathfrak{M}$, se dice que X es un *espacio medible.*

Por ejemplo, podemos tomar $X = R^p$, $\mathfrak{M}$ la colección de todos los subconjuntos medibles de Lebesgue de R^p, y μ la medida de Lebesgue.

O bien, X el conjunto de todos los enteros positivos, $\mathfrak{M}$ la colección de todos los subconjuntos de X, y $\mu(E)$ el número de elementos de E.

Otro ejemplo lo proporciona la teoría de probabilidad, donde pueden considerarse los sucesos como conjuntos y la probabilidad de que se produzcan los sucesos, es una función de conjuntos aditiva (o aditiva en forma numerable).

En los párrafos siguientes trataremos siempre sobre espacios medibles. Se recalca el hecho de que la teoría de la integración que estudiaremos pronto no será más sencilla en ningún aspecto si sacrificásemos la generalidad que hemos alcanzado y nos limitásemos a la medida de Lebesgue, es decir, a un intervalo de la recta real. De hecho, la característica esencial de la teoría se deduce con mucha más claridad en el caso más general, en el que se ve que todo depende solamente de la aditividad numerable de μ en un σ-anillo.

Será conveniente introducir la notación

$$\{x \mid P\} \tag{41}$$

para el conjunto de todos los elementos x que tienen la propiedad P.

FUNCIONES MEDIBLES

11.13 Definición Sea f una función definida en el espacio medible X, con valores en el sistema ampliado de los números reales. Se dice que la función f es *medible* si el conjunto

(42) $$\{x \mid f(x) > a\}$$

es medible para todo número real a.

11.14 Ejemplo Si $X = R^p$ y $\mathfrak{M} = \mathfrak{M}(\mu)$ tal como se define en la Definición 11.9, toda f continua es medible, pues entonces (42) es un conjunto abierto.

11.15 Teorema *Cada una de las cuatro condiciones siguientes implica las otras tres:*

(43) $\{x \mid f(x) > a\}$ *es medible para todo a real.*

(44) $\{x \mid f(x) \geq a\}$ *es medible para todo a real.*

(45) $\{x \mid f(x) < a\}$ *es medible para todo a real.*

(46) $\{x \mid f(x) \leq a\}$ *es medible para todo a real.*

Demostración Las relaciones

$$\{x \mid f(x) \geq a\} = \bigcap_{n=1}^{\infty} \left\{x \mid f(x) > a - \frac{1}{n}\right\},$$

$$\{x \mid f(x) < a\} = X - \{x \mid f(x) \geq a\},$$

$$\{x \mid f(x) \leq a\} = \bigcap_{n=1}^{\infty} \left\{x \mid f(x) < a + \frac{1}{n}\right\},$$

$$\{x \mid f(x) > a\} = X - \{x \mid f(x) \leq a\}$$

demuestra sucesivamente que (43) implica (44), (44) implica (45), (45) implica (46) y (46) implica (43).

Por tanto, puede usarse cualquiera de estas relaciones en lugar de (42) para definir la mensurabilidad.

11.16 Teorema *Si f es medible, $|f|$ es medible.*

Demostración

$$\{x \mid |f(x)| < a\} = \{x \mid f(x) < a\} \cap \{x \mid f(x) > -a\}.$$

11.17 Teorema *Sea $\{f_n\}$ una sucesión de funciones medibles. Para $x \in X$, pongamos*

$$g(x) = \sup f_n(x) \qquad (n = 1, 2, 3, \ldots),$$

$$h(x) = \limsup_{n \to \infty} f_n(x).$$

Entonces, g y h son medibles.

Esto mismo es cierto, naturalmente, para el inf y el lím inf.

Demostración

$$\{x \mid g(x) > a\} = \bigcup_{n=1}^{\infty} \{x \mid f_n(x) > a\},$$
$$h(x) = \inf g_m(x),$$

donde $g_m(x) = \sup f_n(x)$ $(n \geq m)$.

Corolarios

(*a*) *Si f y g son medibles, entonces* máx(f,g) *y* mín(f,g) *son medibles. Si*

(47) $$f^+ = \text{máx}(f, 0), \qquad f^- = -\text{mín}(f, 0),$$

se deduce, en particular que f^+ y f^- son medibles.
(*b*) *El límite de una sucesión convergente de funciones medibles es medible.*

11.18 Teorema *Sean f y g funciones de valores reales medibles definidos en X, sea F real y continua en R^2, y hagamos*

$$h(x) = F(f(x), g(x)) \qquad (x \in X).$$

Entonces, h es medible
En particular, f + g y fg son medibles.

Demostración Sea

$$G_a = \{(u, v) \mid F(u, v) > a\}.$$

G_a es un subconjunto abierto de R^2, y podemos escribir

$$G_a = \bigcup_{n=1}^{\infty} I_n,$$

donde $\{I_n\}$ es una sucesión de intervalos abiertos:

$$I_n = \{(u, v) \mid a_n < u < b_n, c_n < v < d_n\}.$$

Como

$$\{x \mid a_n < f(x) < b_n\} = \{x \mid f(x) > a_n\} \cap \{x \mid f(x) < b_n\}$$

es medible, se deduce que el conjunto

$$\{x \mid (f(x), g(x)) \in I_n\} = \{x \mid a_n < f(x) < b_n\} \cap \{x \mid c_n < g(x) < d_n\}$$

es medible. Por tanto, también sucede lo mismo con

$$\{x \mid h(x) > a\} = \{x \mid (f(x), g(x)) \in G_a\} = \bigcup_{n=1}^{\infty} \{x \mid (f(x), g(x)) \in I_n\}.$$

Recapitulando, podemos decir que todas las operaciones ordinarias del análisis, incluyendo las de límites, cuando se aplican a funciones medibles, producen funciones medibles; en otras palabras, todas las funciones con las que nos encontramos de ordinario, son medibles.

Que, sin embargo, esto no es más que una afirmación aproximada, se demuestra con el siguiente ejemplo (basado en la medida de Lebesgue, en la recta real): Si $h(x) = f(g(x))$, donde f es medible y g continua, h no es necesariamente medible. (Para detalles, véase McShane, p. 241.)

El lector habrá observado que no hemos mencionado la medida en el estudio de las funciones medibles. De hecho la clase de las funciones medibles en X depende solo de σ-anillo (utilizando la notación de la Definición 11.12). Por ejemplo, podemos hablar de funciones medibles según Borel en R^p, esto es, funciones f para las que

$$\{x \mid f(x) > a\}$$

es siempre un conjunto de Borel, sin referencia a ninguna medida particular.

FUNCIONES SIMPLES

11.19 Definición Sea s una función de valores reales definida en X. Si el rango de s es finito, decimos que s es una *función simple.*

Sea $E \subset X$, y pongamos

$$K_E(x) = \begin{cases} 1 & (x \in E), \\ 0 & (x \notin E). \end{cases} \tag{48}$$

a K_E se llama *función característica* de E.

Supongamos que el rango de s consta de los números distintos $c_1, \ldots, c_n$. Sea

$$E_i = \{x \mid s(x) = c_i\} \qquad (i = 1, \ldots, n).$$

Entonces

(49)
$$s = \sum_{n=1}^{n} c_i K_{E_i},$$

esto es, toda función simple es una combinación lineal finita de funciones características. Es evidente que s es medible si, y solo si los conjuntos $E_1, \ldots, E_n$ son medibles.

Es de interés que toda función puede aproximarse por funciones simples:

11.20 Teorema *Sea f una función real en X. Existe una sucesión $\{s_p\}$ de funciones simples tales que $s_n(x) \to f(x)$ si $n \to \infty$,* para todo $x \in X$. *Si f es medible, puede elegirse $\{s_n\}$ de modo que sea una sucesión de funciones medibles. Si $f \geq 0$ $\{s_n\}$ puede elegirse tal que sea una sucesión monótona creciente.*

Demostración Si $f \geq 0$, definiremos

$$E_{ni} = \left\{x \left| \frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n} \right.\right\}, \qquad F_n = \{x | f(x) \geq n\}$$

para $n = 1,2,3,\ldots$, $i = 1,2,\ldots, n2^n$. Pongamos

(50)
$$s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} K_{E_{ni}} + nK_{F_n}.$$

En el caso general, $f = f^+ - f^-$ y aplicaremos la construcción precedente a f^+ y a f^-.

Puede observarse que la sucesión $\{s_n\}$ dada por (50) converge uniformemente hacia f si f es acotada.

INTEGRACIÓN

Definiremos la integración en un espacio medible X, en el que $\mathfrak{M}$ es el σ-anillo de conjuntos medibles, y μ es la medida. El lector que quiera considerar un caso más concreto puede pensar en X como la recta real, o un intervalo, y en μ como la medida m de Lebesgue.

11.21 Definición Supongamos que

(51)
$$s(x) = \sum_{i=1}^{n} c_i K_{E_i}(x) \qquad (x \in X, c_i > 0)$$

es medible, y que $E \in \mathfrak{M}$. Definiremos

(52) $$I_E(s) = \sum_{i=1}^{n} c_i \mu(E \cap E_i).$$

Si f es medible y no negativa, definiremos

(53) $$\int_E f\, d\mu = \sup I_E(s),$$

donde se ha tomado el sup sobre todas las funciones simples s tales que $0 \leq s \leq f$.

Al primer miembro de (53) se le llama *integral de Lebesgue* de f, respecto a la medida μ, sobre el conjunto E. Se observará que la integral puede tener el valor $+\infty$.

Se comprobará fácilmente que

(54) $$\int_E s\, d\mu = I_E(s)$$

para toda función medible simple no negativa s.

11.22 Definición Sea f medible, y consideremos las dos integrales

(55) $$\int_E f^+\, d\mu, \qquad \int_E f^-\, d\mu,$$

donde f^+ y f^- están definidas como en (47).

Si una al menos de las integrales (55) es finita, definiremos

(56) $$\int_E f\, d\mu = \int_E f^+\, d\mu - \int_E f^- d\mu.$$

Si las dos integrales de (55) son finitas, (56) es finita, y diremos que f es *integrable* (o *sumable*) en E en el sentido de Lebesgue, con respecto a μ; y escribiremos $f \in \mathscr{L}(\mu)$ en E. Si $\mu = m$, la notación habitual es: $f \in \mathscr{L}$ en E.

Esta terminología puede resultar un poco confusa: si (56) es $+\infty$ o $-\infty$, la integral de f sobre E está definida, aunque f no es integrable en el sentido anterior de la palabra; f es integrable en E solamente si su integral sobre E es finita.

Nos interesaremos principalmente por las funciones integrables aunque en algunos casos es conveniente tratar el caso más general.

11.23 Observaciones Las propiedades siguientes son evidentes:

(*a*) Si f es medible y acotada en E, y si $\mu(E) < +\infty$, entonces $f \in \mathscr{L}(\mu)$ en E.

(*b*) Si $a \leq f(x) \leq b$ para $x \in E$, y $\mu(E) < +\infty$, es

$$a\mu(E) \leq \int_E f\,d\mu \leq b\mu(E).$$

(*c*) Si f y $g \in \mathscr{L}(\mu)$ en E, y si $f(x) \leq g(x)$ para $x \in E$

$$\int_E f\,d\mu \leq \int_E g\,d\mu.$$

(*d*) Si $f \in \mathscr{L}(\mu)$ en E, entonces $cf \in \mathscr{L}(\mu)$ en E, para toda constante finita c, y

$$\int_E cf\,d\mu = c\int_E f\,d\mu.$$

(*e*) Si $\mu(E) = 0$, y f es medible,

$$\int_E f\,d\mu = 0.$$

(*f*) Si $f \in \mathscr{L}(\mu)$ en E, $A \in \mathfrak{M}$, y $A \subset E$, entonces $f \in \mathscr{L}(\mu)$ en A.

11.24 Teorema

(*a*) *Supongamos que f es medible y no negativa en X. Para $A \in \mathfrak{M}$, definamos*

(57) $$\phi(A) = \int_A f\,d\mu.$$

Entonces, ϕ es aditiva numerable en $\mathfrak{M}$.
(*b*) *La misma conclusión se cumple si $f \in \mathscr{L}(\mu)$ en X.*

Demostración Es evidente que (*b*) se deduce de (*a*) si escribimos $f = f^+ - f^-$ y aplicamos (*a*) a f^+ y a f^-.

Para demostrar (*a*), tenemos que probar que

(58) $$\phi(A) = \sum_{n=1}^{\infty} \phi(A_n)$$

Si $A_n \in \mathfrak{M}\,(n = 1,2,3,\ldots)$, si $A_i \cap A_j = 0$ para $i \neq j$, y si $A = \bigcup_1^\infty A_n$.

Si f es una función característica, la aditividad numerable de ϕ es precisamente la misma que la aditividad numerable de μ, pues

$$\int_A K_E \, d\mu = \mu(A \cap E).$$

Si f es simple, es de la forma (51) y la conclusión se sigue cumpliendo.

En el caso general, tenemos para cada función simple medible s tal que $0 \leq s \leq f$.

$$\int_A s \, d\mu = \sum_{n=1}^{\infty} \int_{A_n} s \, d\mu \leq \sum_{n=1}^{\infty} \phi(A_n).$$

Por tanto, por (53),

(59) $$\phi(A) \leq \sum_{n=1}^{\infty} \phi(A_n).$$

Ahora, si $\phi(A_n) = +\infty$ para cualquier n, (58) es trivial, porque $\phi(A) \geq \phi(A_n)$. Supongamos que $\phi(A_n) < +\infty$ para todo n.

Dado $\varepsilon > 0$, podemos elegir una función medible s tal que $0 \leq s \leq f$, y que

(60) $$\int_{A_1} s \, d\mu \geq \int_{A_1} f \, d\mu - \varepsilon, \qquad \int_{A_2} s \, d\mu \geq \int_{A_2} f \, d\mu - \varepsilon.$$

Por lo tanto,

$$\phi(A_1 \cup A_2) \geq \int_{A_1 \cup A_2} s \, d\mu = \int_{A_1} s \, d\mu + \int_{A_2} s \, d\mu \geq \phi(A_1) + \phi(A_2) - 2\varepsilon,$$

de modo que

$$\phi(A_1 \cup A_2) \geq \phi(A_1) + \phi(A_2).$$

Se deduce que tenemos, para todo n,

(61) $$\phi(A_1 \cup \cdots \cup A_n) \geq \phi(A_1) + \cdots + \phi(A_n).$$

Como $A \supset A_1 \cup \cdots \cup A_n$, (61) implica

(62) $$\phi(A) \geq \sum_{n=1}^{\infty} \phi(A_n),$$

y de (59) y (62) se deduce (58).

Corolario *Si* $A \in \mathfrak{M}$, $B \subset A$, *y* $\mu(A - B) = 0$, *será*

$$\int_A f\,d\mu = \int_B f\,d\mu.$$

Como $A = B \cup (A - B)$, esta expresión se deduce de la observación 11.23(*e*).

11.25 Observaciones El corolario anterior demuestra que los conjuntos de medida nula son despreciables en la integración.

Escribamos $f \sim g$ en E si el conjunto

$$\{x \mid f(x) \neq g(x)\} \cap E$$

tiene medida nula.

Entonces, $f \sim f$; $f \sim g$, implica $g \sim f$, y $f \sim g$, $g \sim h$ implica $f \sim h$. Esto es, la relación $\sim$ es una relación de equivalencia.

Si $f \sim g$ en E, tenemos como se ve fácilmente,

$$\int_A f\,d\mu = \int_A g\,d\mu,$$

siempre que existan las integrales, para todo subconjunto medible A de E.

Si se cumple una propiedad P para todo $x \in E - A$, y si $\mu(A) = 0$, se acostumbra a decir que P se cumple para casi todo $x \in E$, o que P se cumple casi en todas partes en E. (Este concepto «casi en todas partes» depende, ciertamente, de la medida particular que se considere. En el texto, si no se dice nada en contrario, se refiere de ordinario a la medida de Lebesgue.)

Si $f \in \mathscr{L}(\mu)$ en E, es evidente que $f(x)$ debe ser finita en casi todas partes en E. Sin embargo, en la mayoría de los casos no perderemos ninguna generalidad si suponemos que las funciones dadas tienen valores finitos desde el principio.

11.26 Teorema *Si* $f \in \mathscr{L}(\mu)$ *en* E, *entonces* $|f| \in \mathscr{L}(\mu)$ *en* E, *y*

$$\left|\int_E f\,d\mu\right| \leq \int_E |f|\,d\mu. \tag{63}$$

Demostración Escribamos $E = A \cup B$, siendo $f(x) \geq 0$ en A y $f(x) < 0$ en B. Por el Teorema 11.24.

$$\int_E |f|\,d\mu = \int_A |f|\,d\mu + \int_B |f|\,d\mu = \int_A f^+\,d\mu + \int_B f^-\,d\mu < +\infty,$$

de modo que $|f| \in \mathscr{L}(\mu)$: Como $f \leq |f|$ y $-f \leq |f|$, vemos que

$$\int_E f\,d\mu \leq \int_E |f|\,d\mu, \qquad -\int_E f\,d\mu \leq \int_E |f|\,d\mu,$$

y se deduce (63).

Como la integrabilidad de f implica la de $|f|$, de la integral de Lebesgue se dice a menudo que es una integral absolutamente convergente. Es ciertamente posible definir integrales no absolutamente convergentes, y en la resolución de algunos problemas es esencial hacerlo; pero estas integrales carecen de alguna de las propiedades más útiles de la integral de Lebesgue y juegan un papel algo menos importante en el análisis.

11.27 Teorema *Supongamos que f es medible en E, $|f| \leq g$, y $g \in \mathscr{L}(\mu)$ en E. Entonces $f \in \mathscr{L}(\mu)$ en E.*

Demostración Tenemos que $f^+ \leq g$ y $f^- \leq g$.

11.28 Teorema de la convergencia monótona de Lebesgue *Supongamos que $E \in \mathfrak{M}$. Sea $\{f_n\}$ una sucesión de funciones medibles, tales que*

(64) $$0 \leq f_1(x) \leq f_2(x) \leq \cdots \qquad (x \in E).$$

Sea f definida por

(65) $$f_n(x) \to f(x) \qquad (x \in E)$$

cuando $n \to \infty$. Entonces

(66) $$\int_E f_n\,d\mu \to \int_E f\,d\mu \qquad (n \to \infty).$$

Demostración Por (64), es claro que, cuando $n \to \infty$,

(67) $$\int_E f_n\,d\mu \to \alpha$$

para algún α; y como $\int f_n \leq \int f$, tenemos

(68) $$\alpha \leq \int_E f\,d\mu.$$

Elijamos c, tal que $0 < c < 1$, y sea s una función simple medible, tal que $0 \leq s \leq f$. Pongamos

$$E_n = \{x \mid f_n(x) \geq cs(x)\} \qquad (n = 1, 2, 3, \ldots).$$

Por (64), $E_1 \subset E_2 \subset E_3 \subset, \ldots,$ y por (65),

(69) $$E = \bigcup_{n=1}^{\infty} E_n.$$

Para todo n,

(70) $$\int_E f_n \, d\mu \geq \int_{E_n} f_n \, d\mu \geq c \int_{E_n} s \, d\mu.$$

Hagamos $n \to \infty$ en (70). Como la integral es una función de conjuntos aditiva numerable (Teorema 11.24); (69) demuestra que podemos aplicar el Teorema 11.3 a la última integral de (70) y obtenemos

(71) $$\alpha \geq c \int_E s \, d\mu.$$

Haciendo $c \to 1$, vemos que

$$\alpha \geq \int_E s \, d\mu,$$

y (53) implica

(72) $$\alpha \geq \int_E f \, d\mu.$$

Deduciéndose el teorema de (67), (68) y (72).

11.29 Teorema *Supongamos que $f = f_1 + f_2$, donde $f_i \in \mathscr{L}(\mu)$ sobre E ($i = 1,2$). Entonces $f \in \mathscr{L}(\mu)$ en E, y*

(73) $$\int_E f \, d\mu = \int_E f_1 \, d\mu + \int_E f_2 \, d\mu.$$

Demostración Supongamos, primero, que $f_1 \geq 0$, $f_2 \geq 0$. Si f_1 y f_2 son simples, de (52) y (54) se deduce inmediatamente (73). En caso contrario, elijamos sucesiones monótonas crecientes $\{s_n'\}$, $\{s_n''\}$ de funciones simples medibles no negativas que convergen hacia f_1 y f_2. El Teorema 11.20 demuestra que es posible. Pongamos $s_n = s_n' + s_n''$. Será

$$\int_E s_n \, d\mu = \int_E s_n' \, d\mu + \int_E s_n'' \, d\mu,$$

y se deduce (73) si hacemos $n \to \infty$ y recurrimos al Teorema 11.28.

Continuando, supongamos que $f_1 \geq 0$, $f_2 \leq 0$. Pongamos

$$A = \{x \mid f(x) \geq 0\}, \qquad B = \{x \mid f(x) < 0\}.$$

Entonces f, f_1 y $-f_2$ son no negativas en A. Por tanto,

(74)
$$\int_A f_1 d\mu = \int_A f\, d\mu + \int_A (-f_2)\, d\mu = \int_A f\, d\mu - \int_A f_2\, d\mu.$$

De igual modo, $-f, f_1$ y $-f_2$ son no negativas en B, de modo que

$$\int_B (-f_2)\, d\mu = \int_B f_1\, d\mu + \int_B (-f)\, d\mu,$$

o

(75)
$$\int_B f_1\, d\mu = \int_B f\, d\mu - \int_B f_2\, d\mu,$$

y se deduce (73) si sumamos (74) y (75).

En el caso general, puede descomponerse E en cuatro conjuntos E_i, en cada uno de los cuales tienen signo constante $f_1(x)$ y $f_2(x)$. Los dos casos que ya hemos demostrado implican

$$\int_{E_i} f\, d\mu = \int_{E_i} f_1\, d\mu + \int_{E_i} f_2\, d\mu \qquad (i = 1, 2, 3, 4),$$

y se deduce (73) sumando esas cuatro ecuaciones.

Ahora estamos en condiciones de volver a enunciar el Teorema 11.28 para las series:

11.30 Teorema *Supongamos que $E \in \mathfrak{M}$. Si $\{f_n\}$ es una sucesión de funciones medibles no negativas y*

(76)
$$f(x) = \sum_{n=1}^{\infty} f_n(x) \qquad (x \in E),$$

entonces

$$\int_E f\, d\mu = \sum_{n=1}^{\infty} \int_E f_n\, d\mu.$$

Demostración Las sumas parciales de (76) forman una sucesión monótona creciente.

11.31 Teorema de Fatou *Supongamos que $E \in \mathfrak{M}$. Si $\{f_n\}$ es una sucesión de funciones medibles no negativas, y*

$$f(x) = \lim_{n\to\infty} \inf f_n(x) \qquad (x \in E),$$

entonces

(77) $$\int_E f\, d\mu \le \lim_{n\to\infty} \inf \int_E f_n\, d\mu.$$

En (77) puede darse la desigualdad. En el Ejercicio 5 se da un ejemplo.

Demostración Para $n = 1,2,3,\ldots$ y $x \in E$, hagamos

$$g_n(x) = \inf f_i(x) \qquad (i \ge n).$$

Entonces g_n es medible en E, y

(78) $$0 \le g_1(x) \le g_2(x) \le \cdots,$$

(79) $$g_n(x) \le f_n(x),$$

(80) $$g_n(x) \to f(x) \qquad (n \to \infty).$$

Por (78); (80) y el Te orema 11.28,

(81) $$\int_E g_n\, d\mu \to \int_E f\, d\mu,$$

de modo que de (79) y (81) se deduce (77).

11.32 Teorema de la convergencia dominada de Lebesgue *Supongamos que $E \in \mathfrak{M}$. Sea $\{f_n\}$ una sucesión de funciones medibles, tales que*

(82) $$f_n(x) \to f(x) \qquad (x \in E)$$

Cuando $n \to \infty$. si existe una función $g \in \mathscr{L}(\mu)$ en E, tal que

(83) $$|f_n(x)| \le g(x) \qquad (n = 1, 2, 3, \ldots, x \in E),$$

entonces

(84) $$\lim_{n\to\infty} \int_E f_n\, d\mu = \int_E f\, d\mu.$$

A consecuencia de (83), se dice que $\{f_n\}$ está dominada por g, y hablaremos de convergencia dominada. Por la Observación 11.25, la conclusión es la misma si se cumple (82) en casi todas partes en E.

Demostración Primeramente (83) y el Teorema 11.27 implican que $f_n \in \mathscr{L}(\mu)$ y $f \in \mathscr{L}(\mu)$ en E.

Como $f_n + g \geq 0$, el teorema de Fatou demuestra que

$$\int_E (f+g)\,d\mu \leq \liminf_{n\to\infty} \int_E (f_n+g)\,d\mu,$$

o

(85) $$\int_E f\,d\mu \leq \liminf_{n\to\infty} \int_E f_n\,d\mu.$$

Como $g - f_n \geq 0$, vemos igualmente que

$$\int_E (g-f)\,d\mu \leq \liminf_{n\to\infty} \int_E (g-f_n)\,d\mu,$$

de modo que

$$-\int_E f\,d\mu \leq \liminf_{n\to\infty} \left[-\int_E f_n\,d\mu\right],$$

que es lo mismo que

(86) $$\int_E f\,d\mu \geq \limsup_{n\to\infty} \int_E f\,d\mu.$$

La existencia del límite en (84) y la igualdad planteada en (84) se deducen, ahora, de (85) y (86).

Corolario *Si $\mu(E) < +\infty$, $\{f_n\}$, es uniformemente acotada en E, y $f_n(x) \to f(x)$, en E se cumple* (84).

Una sucesión convergente uniformemente acotada, se llama frecuentemente acotadamente convergente.

COMPARACIÓN CON LA INTEGRAL DE RIEMANN

El próximo teorema demuestra que toda función que es integrable, según Riemann, en un intervalo es, también integrable, según Lebesgue, y que las

funciones integrables según Riemann están sujetas a condiciones de continuidad más rigurosas. Aparte del hecho de que la teoría de Lebesgue nos permite, por ello, integrar una clase mucho más amplia de funciones, su mayor ventaja estriba quizá en la facilidad con que se manejan muchas operaciones de límites; desde este punto de vista, los teoremas de convergencia de Lebesgue pueden considerarse como el corazón de la teoría de Lebesgue.

Una de las dificultades que se encuentra en la teoría de Riemann, es que los límites de las funciones integrables según Riemann (o incluso de las funciones continuas) pueden no ser integrables según Riemann. Esta dificultad queda ahora casi eliminada, pues, los límites de funciones medibles son siempre medible.

Sea el espacio de medida X el intervalo $[a,b]$ de la recta real, con $\mu = m$ (medida de Lebesgue), y $\mathfrak{M}$ la familia de subconjuntos de $[a,b]$ medibles, según Lebesgue. En lugar de

$$\int_X f\,dm$$

se acostumbra usar la notación ordinaria

$$\int_a^b f\,dx$$

para la integral de Lebesgue de f sobre $[a,b]$. Para distinguir las integrales de Riemann de las de Lebesgue, representaremos las primeras por

$$\mathscr{R}\int_a^b f\,dx.$$

11.33 Teorema

(a) *Si $f \in \mathscr{R}$ en $[a,b]$, entonces $f \in \mathscr{L}$ en $[a,b]$, y*

(87) $$\int_a^b f\,dx = \mathscr{R}\int_a^b f\,dx.$$

(b) *Supongamos que f es acotada en $[a,b]$. Entonces $f \in \mathscr{R}$ si, y solo si f es continua en casi todas partes en $[a,b]$.*

Demostración Supóngase que f es acotada. Por la Definición 6.1 y el Teorema 6.4 existe una sucesión $\{P_k\}$ de particiones de $[a,b]$ tal que P_{k+1} es un refinamiento de P_k, tal que la distancia entre puntos adyacentes de P_k es menor que $1/k$, y tal que

(88) $$\lim_{k\to\infty} L(P_k, f) = \mathscr{R}\underline{\int} f\,dx, \qquad \lim_{k\to\infty} U(P_k, f) = \mathscr{R}\overline{\int} f\,dx.$$

(En esta demostración, todas las integrales se toman sobre $[a,b]$.)

Si $P_k = \{x_0, x_1, \ldots, x_n\}$, con $x_0 = a, x_n = b$, define

$$U_k(a) = L_k(a) = f(a);$$

poniendo $U_k(x) = M_i$ y $L_k(x) = m_i$ para $x_{i-1} < x \le x_i$, $1 \le i \le n$, y usando la notación que se introdujo en la Definición 6.1. Entonces

(89) $$L(P_k, f) = \int L_k\,dx, \qquad U(P_k, f) = \int U_k\,dx,$$

y

(90) $$L_1(x) \le L_2(x) \le \cdots \le f(x) \le \cdots \le U_2(x) \le U_1(x)$$

para todo $x \in [a,b]$, ya que P_{k+1} refina a P_k. Por (90), existe

(91) $$L(x) = \lim_{k\to\infty} L_k(x), \qquad U(x) = \lim_{k\to\infty} U_k(x).$$

Obsérvese que L y U son funciones medibles acotadas sobre $[a,b]$, y que

(92) $$L(x) \le f(x) \le U(x) \qquad (a \le x \le b),$$

y que por (88), (90) y el teorema de convergencia monótona,

(93) $$\int L\,dx = \mathscr{R}\underline{\int} f\,dx, \qquad \int U\,dx = \mathscr{R}\overline{\int} f\,dx,$$

Hasta aquí, no se ha supuesto nada acerca de f excepto que f es una función real acotada sobre $[a,b]$.

Para completar la demostración, nótese que $f \in \mathscr{R}$ si y solo si sus integrales superior e inferior son iguales; en consecuencia si y solo si

(94) $$\int L\,dx = \int U\,dx;$$

ya que $L \le U$, (94) sucede si y solo si $L(x) = U(x)$ a lo más para $x \in [a,b]$ (Ejercicio 1).

En este caso, (92) implica que

$$L(x) = f(x) = U(x) \tag{95}$$

casi en todas partes sobre $[a,b]$, así que f es medible, y (87) se deduce de (93) y (95).

Además, si x pertenece o no a P_k, se puede ver fácilmente que $U(x) = L(x)$ si y solo si f es continua en x. Como la unión de los conjuntos P_k es numerable, su medida es 0 y se concluye que f es continua casi en todas partes sobre $[a,b]$ si y solo si $L(x) = U(x)$ casi en todas partes, en consecuencia (como se vio anteriormente) si, y solo si $f \in \mathscr{R}$.

Esto completa la demostración.

La relación ordinaria entre integración y diferenciación se traslada en alto grado a la teoría de Lebesgue. Si $f \in \mathscr{L}$ en $[a,b]$, y

$$F(x) = \int_a^x f\,dt \qquad (a \leq x \leq b), \tag{96}$$

será $F'(x) = f(x)$ en casi todas partes en $[a,b]$.

Inversamente, si F es diferenciable en todo punto de $[a,b]$ (aquí no sirve ya «casi en todas partes») y si $F' \in \mathscr{L}$ en $[a,b]$, será

$$F(x) - F(a) = \int_a^x F'(t)\,dt \qquad (a \leq x \leq b).$$

Para la demostración de estos teoremas, aconsejamos al lector vea alguna de las obras sobre integración citadas en la Bibliografía.

INTEGRACIÓN DE FUNCIONES COMPLEJAS

Supongamos que f es una función compleja definida en un espacio de medida X, y $f = u + iv$, donde u y v son reales. Decimos que f es medible si, y solo si tanto u como v son medibles.

Es fácil comprobar que las sumas y productos de funciones medibles complejas son también medibles. Como

$$|f| = (u^2 + v^2)^{1/2},$$

el Teorema 11.18 demuestra que $|f|$ es medible para toda f medible compleja.

Supongamos que μ es una medida en X, E es un subconjunto medible de X y f una función compleja en X. Decimos que $f \in \mathscr{L}(\mu)$ en E siempre que f sea medible y

$$\int_E |f|\,d\mu < +\infty, \tag{97}$$

y definimos

$$\int_E f\,d\mu = \int_E u\,d\mu + i\int_E v\,d\mu$$

si se cumple (97). Como $|u| \leq |f|$, $|v| \leq |f|$ y $|f| \leq |u| + |v|$, es evidente que se cumple (97) si, y solo si $u \in \mathscr{L}(\mu)$ y $v \in \mathscr{L}(\mu)$ en E.

Ahora pueden extenderse los Teoremas 11.23(a), (d), (e), (f), 11.24(b), 11.26, 11.27, 11.29 y 11.32 a las integrales de Lebesgue de funciones complejas. Las demostraciones son totalmente directas. La del Teorema 11.26 es la única que ofrece algo de interés:

Si $f \in \mathscr{L}(\mu)$ en E, hay un número complejo c, $|c| = 1$, tal que

$$c\int_E f\,d\mu \geq 0.$$

Hagamos $g = cf = u + iv$, u y v reales. Entonces,

$$\left|\int_E f\,d\mu\right| = c\int_E f\,d\mu = \int_E g\,d\mu = \int_E u\,d\mu \leq \int_E |f|\,d\mu.$$

La tercera de las igualdades anteriores se cumplen, pues las precedentes demuestran que $\int_g d\mu$ es real.

FUNCIONES DE CLASE $\mathscr{L}^2$

Como aplicación de la teoría de Lebesgue, ampliaremos el teorema de Parseval (que demostramos solo para funciones continuas en el Cap. 8) y demostraremos el teorema de Riesz-Fischer para conjuntos ortonormales de funciones.

11.34 Definición Sea X un espacio medible. Decimos que una función compleja $f \in \mathscr{L}^2(\mu)$ en X, si f es medible y si

$$\int_X |f|^2\,d\mu < +\infty.$$

Si μ es una medida de Lebesgue decimos que $f \in \mathscr{L}^2$. Para $f \in \mathscr{L}^2(\mu)$ (desde ahora omitiremos la expresión «en X») definimos

$$\|f\| = \left\{\int_X |f|^2\,d\mu\right\}^{1/2}$$

y llamamos a $\|f\|$ la norma $\mathscr{L}^2(\mu)$ de f.

11.35 Teorema *Supongamos que $f \in \mathscr{L}^2(\mu)$ y $g \in \mathscr{L}^2(\mu)$, Entonces $fg \in \mathscr{L}(\mu)$, y*

(98) $$\int_X |fg|\,d\mu \leq \|f\|\,\|g\|.$$

Esta es la desigualdad de Schwarz que ya hemos encontrado para las series y para las integrales de Riemann. De la desigualdad se deduce que

$$0 \leq \int_X (|f| + \lambda|g|)^2 \, d\mu = \|f\|^2 + 2\lambda \int_X |fg| \, d\mu + \lambda^2 \|g\|^2,$$

que es cierta para todo λ real.

11.36 Teorema *Si $f \in \mathscr{L}^2(\mu)$ y $g \in \mathscr{L}^2(\mu)$, entonces $f + g \in \mathscr{L}^2(\mu)$, y*

$$\|f + g\| \leq \|f\| + \|g\|.$$

Demostración *La desigualdad de Schwarz muestra que*

$$\begin{aligned} \|f+g\|^2 &= \int |f|^2 + \int f\bar{g} + \int \bar{f}g + \int |g|^2 \\ &\leq \|f\|^2 + 2\|f\| \, \|g\| + \|g\|^2 \\ &= (\|f\| + \|g\|)^2. \end{aligned}$$

11.37 Observación Si definimos la distancia entre dos funciones f y g en $\mathscr{L}^2(\mu)$ por $\|f - g\|$, vemos que se satisfacen las condiciones de la Definición 2.17, excepto por el hecho de que $\|f - g\| = 0$ no implica que $f(x) = g(x)$ para todo x, sino solo para casi todo x. Así, si consideramos idénticas las funciones que difieren solo en un conjunto de medida cero, $\mathscr{L}^2(\mu)$ es un espacio métrico.

Consideraremos ahora $\mathscr{L}^2$ en un intervalo de la recta real, con respecto a la medida de Lebesgue.

11.38 Teorema *Las funciones continuas forman un subconjunto denso de $\mathscr{L}^2$ en $[a,b]$.*

Más explícitamente, esto significa que para cualquier $f \in \mathscr{L}^2$ en $[a,b]$, y cualquier $\varepsilon > 0$, hay una función g, continua en $[a,b]$, tal que

$$\|f - g\| = \left\{\int_a^b (f-g)^2 \, dx\right\}^{1/2} < \varepsilon.$$

Demostración Diremos que f es aproximada en $\mathscr{L}^2$ por una sucesión $\{g_n\}$ si $\|f - g_n\| \to 0$ cuando $n \to \infty$.

Sea A un subconjunto cerrado de $[a,b]$ y K_A su función característica. Pongamos

$$t(x) = \inf |x - y| \qquad (y \in A)$$

y
$$g_n(x) = \frac{1}{1 + nt(x)} \qquad (n = 1, 2, 3, \ldots).$$

Entonces g_n es continua en $[a,b]$, $g_n(x) = 1$ en A, y $g_n(x) \to 0$ en B, siendo $B = [a,b] - A$. Por tanto,

$$\|g_n - K_A\| = \left\{\int_B g_n^2\, dx\right\}^{1/2} \to 0$$

por el Teorema 11.32. Así pues, las funciones características de conjuntos cerrados pueden ser aproximadas en $\mathscr{L}^2$ por funciones continuas.

Por (39) se ve que también es cierto para la función característica de cualquier conjunto medible, y por consiguiente, también para las funciones medibles simples.

Si $f \geq 0$ y $f \in \mathscr{L}^2$, sea $\{s_n\}$ una sucesión monótona creciente de funciones medibles simples no negativas tales que $s_n(x) \to f(x)$. Como $|f - s_n|^2 \leq f^2$, el Teorema 11.32 demuestra que $\|f - s_n\| \to 0$.

De aquí se deduce el caso general.

11.39 Definición Decimos que una sucesión de funciones complejas $\{\phi_n\}$ es un conjunto *ortonormal* de funciones en un espacio medible X, si

$$\int_X \phi_n \bar{\phi}_m\, d\mu = \begin{cases} 0 & (n \neq m), \\ 1 & (n = m). \end{cases}$$

En particular, debemos tener $\phi_n \in \mathscr{L}^2(\mu)$. Si $f \in \mathscr{L}^2(\mu)$, y si

$$c_n = \int_X f\bar{\phi}_n\, d\mu \qquad (n = 1, 2, 3, \ldots),$$

escribiremos

$$f \sim \sum_{n=1}^{\infty} c_n \phi_n,$$

como en la Definición 8.10.

La definición de serie de Fourier trigonométrica se amplía del mismo modo a $\mathscr{L}^2$ (o incluso a $\mathscr{L}$) en $[-\pi,\pi]$. Los Teoremas 8.11 y 8.12 (la desigualdad de Bessel) se cumple para cualquier $f \in \mathscr{L}^2(\mu)$. Las demostraciones son iguales, palabra por palabra.

Ahora podemos demostrar el teorema de Parseval.

11.40 Teorema *Supongamos que*

(99)
$$f(x) \sim \sum_{-\infty}^{\infty} c_n e^{inx},$$

donde $f \in \mathscr{L}^2$ en $[-\pi,\pi]$. *Sea* s_n *la n-ésima suma parcial de* (99). *Entonces,*

(100) $$\lim_{n\to\infty} \|f - s_n\| = 0,$$

(101) $$\sum_{-\infty}^{\infty} |c_n|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f|^2\, dx.$$

Demostración Sea $\varepsilon >$ dado. Por el Teorema 11.38, hay una función continua g tal que

$$\|f - g\| < \frac{\varepsilon}{2}.$$

Además, es fácil ver que podemos hacer que $g(\pi) = g(-\pi)$. Entonces puede prolongarse g en una función continua periódica. Por el Teorema 8.16, hay un polinomio trigonométrico T, de grado N, tal que

$$\|g - T\| < \frac{\varepsilon}{2}.$$

Por tanto, por el Teorema 8.11 (ampliado a $\mathscr{L}^2$), $n \geq N$ implica

$$\|s_n - f\| \leq \|T - f\| < \varepsilon,$$

y se deduce (100). La ecuación (102) se deduce de (100) como en la demostración del Teorema 8.16.

Corolario *Si* $f \in \mathscr{L}^2$ en $[-\pi,\pi]$ *y si*

$$\int_{-\pi}^{\pi} f(x)e^{-inx}\, dx = 0 \qquad (n = 0, \pm 1, \pm 2, \ldots),$$

entonces $\|f\| = 0$.

Así, si dos funciones en $\mathscr{L}^2$ tienen la misma serie de Fourier, difieren a lo sumo en un conjunto de medida cero.

11.41 Definición Sean f y $f_n \in \mathscr{L}^2(\mu)$ $(n = 1,2,3,\ldots)$. Decimos que $\{f_n\}$ converge hacia f en $\mathscr{L}^2(\mu)$ si $\|f_n - f\| \to 0$. Decimos que $\{f_n\}$ es una sucesión de Cauchy en $\mathscr{L}^2(\mu)$ si para cada $\epsilon > 0$ hay un entero N tal que $n \geq N$, $m \geq N$ implica $\|f_n - f_m\| \leq \varepsilon$.

11.42 Teorema *Si* $\{f_n\}$ *es una sucesión de Cauchy en* $\mathscr{L}^2(\mu)$, *existe una función* $f \in \mathscr{L}^2(\mu)$ *tal que* $\{f_n\}$ *converge hacia* f *en* $\mathscr{L}^2(\mu)$.

Esto expresa, en otras palabras, que $\mathscr{L}^2(\mu)$ es un espacio métrico *completo.*

Demostración Como $\{f_n\}$ es una sucesión de Cauchy, podemos hallar una sucesión $\{n_k\}$, $k = 1,2,3,\ldots$, tal que

$$\|f_{n_k} - f_{n_{k+1}}\| < \frac{1}{2^k} \qquad (k = 1, 2, 3, \ldots).$$

Elijamos una función $g \in \mathscr{L}^2(\mu)$. Por la desigualdad de Schwarz,

$$\int_X |g(f_{n_k} - f_{n_{k+1}})| \, d\mu \leq \frac{\|g\|}{2^k}.$$

Por tanto,

(102)
$$\sum_{k=1}^{\infty} \int_X |g(f_{n_k} - f_{n_{k+1}})| \, d\mu \leq \|g\|.$$

Por el Teorema 11.30, podemos permutar la suma y la integración en (102). Se deduce que

(103)
$$|g(x)| \sum_{k=1}^{\infty} |f_{n_k}(x) - f_{n_{k+1}}(x)| < +\infty$$

en casi todas partes en X. Por tanto,

(104)
$$\sum_{k=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)| < +\infty$$

en casi todas partes en X. Porque si la serie de (104) fuera divergente en un conjunto E de medida positiva, podríamos tomar $g(x)$ no nula en un subconjunto de E de medida positiva, obteniendo así una contradicción de (103).

Como la k-ésima suma parcial de la serie

$$\sum_{k=1}^{\infty} (f_{n_{k+1}}(x) - f_{n_k}(x)),$$

que converge en casi todas partes en X, es

$$f_{n_{k+1}}(x) - f_{n_1}(x),$$

vemos que la ecuación

$$f(x) = \lim_{k \to \infty} f_{n_k}(x)$$

define $f(x)$ para casi todo $x \in X$ y no importa cómo definamos $f(x)$ en los restantes puntos de X.

Demostraremos ahora que esta función f tiene las propiedades deseadas. Sea $\varepsilon > 0$ dado, y elijamos N como se indica en la Definición 11.41. Si $n_k > N$, el teorema de Fatou demuestra que

$$\|f - f_{n_k}\| \leq \liminf_{i \to \infty} \|f_{n_i} - f_{n_k}\| \leq \varepsilon.$$

Así pues, $f - f_{n_k} \in \mathscr{L}^2(\mu)$, y como $f = (f - f_{n_k}) + f_{n_k}$, vemos que $f \in \mathscr{L}^2(\mu)$. También, como ϵ es arbitrario,

$$\lim_{k \to \infty} \|f - f_{n_k}\| = 0.$$

Finalmente, la desigualdad

(105) $$\|f - f_n\| \leq \|f - f_{n_k}\| + \|f_{n_k} - f_n\|$$

demuestra que $\{f_n\}$ converge hacia f en $\mathscr{L}^2(\mu)$; porque si tomamos n y n_k suficientemente grandes, cada uno de los dos términos del segundo miembro de (105) se puede hacer arbitrariamente pequeño.

11.43 Teorema de Riesz-Fischer *Sea $\{\phi_n\}$ ortonormal en X. Supongamos que* $\Sigma |c_n|^2$ converge, y pongamos $s_n = c_1\phi_1 + \cdots + c_n\phi_n$. *Entonces, existe una función $f \in \mathscr{L}^2(\mu)$ tal que $\{s_n\}$ converge hacia f en $\mathscr{L}^2(\mu)$, y que*

$$f \sim \sum_{n=1}^{\infty} c_n \phi_n.$$

Demostración Para $n > m$,

$$\|s_n - s_m\|^2 = |c_{m+1}|^2 + \cdots + |c_n|^2,$$

de modo que $\{s_n\}$ es una sucesión de Cauchy en $\mathscr{L}^2(\mu)$. Por el Teorema 11.42, hay una función $f \in \mathscr{L}^2(\mu)$ tal que

$$\lim_{n \to \infty} \|f - s_n\| = 0.$$

Ahora, para $n > k$,

$$\int_X f\bar{\phi}_k \, d\mu - c_k = \int_X f\bar{\phi}_k \, d\mu - \int_X s_n \bar{\phi}_k \, d\mu,$$

de modo que

$$\left| \int_X f\bar{\phi}_k \, d\mu - c_k \right| \leq \|f - s_n\| \cdot \|\phi_k\| + \|f - s_n\|.$$

Haciendo $n \to \infty$, vemos que

$$c_k = \int_X f\bar{\phi}_k \, d\mu \qquad (k = 1, 2, 3, \ldots),$$

quedando completa la demostración.

11.44 Definición Se dice que un conjunto ortonormal $\{\phi_n\}$ es *completo* si, para $f \in \mathscr{L}^2(\mu)$, las ecuaciones

$$\int_X f\bar{\phi}_n \, d\mu = 0 \qquad (n = 1, 2, 3, \ldots)$$

implican que $\|f\| = 0$.

En el corolario del Teorema 11.40 dedujimos la completitud del sistema trigonométrico a partir de la ecuación Parseval (101). Inversamente, la ecuación de Parseval se cumple para todo conjunto ortonormal completo:

11.45 Teorema *Sea $\{\phi_n\}$ un conjunto ortonormal completo. Si $f \in \mathscr{L}^2(\mu)$ y si*

(106)
$$f \sim \sum_{n=1}^{\infty} c_n \phi_n,$$

será

(107)
$$\int_X |f|^2 \, d\mu = \sum_{n=1}^{\infty} |c_n|^2.$$

Demostración Por la desigualdad de Bessel, $\Sigma |c_n|^2$ converge. Poniendo

$$s_n = c_1\phi_1 + \cdots + c_n\phi_n,$$

el teorema de Riesz-Fischer demuestra que hay una función $g \in \mathscr{L}^2(\mu)$ tal que

(108)
$$g \sim \sum_{n=1}^{\infty} c_n \phi_n,$$

y que $\|g - s_n\| \to 0$. Por tanto $\|s_n\| \to \|g\|$. Como

$$\|s_n\|^2 = |c_1|^2 + \cdots + |c_n|^2,$$

tenemos

(109)
$$\int_X |g|^2 \, d\mu = \sum_{n=1}^{\infty} |c_n|^2.$$

Ahora, (106), (108) y la completitud de $\{\phi_n\}$, demuestran que $\|f - g\| = 0$, de modo que (109) implica (107).

Combinando los Teoremas 11.43 y 11.45, llegamos a la conclusión muy interesante, de que todo conjunto ortonormal completo induce una correspondencia 1-1 entre las funciones $f \in \mathscr{L}^2(\mu)$. (Considerando idénticas las que son iguales en casi todas partes) por un lado y las sucesiones $\{c_n\}$ para las que converge $\Sigma |c_n|^2$, por otro. La representación

$$f \sim \sum_{n=1}^{\infty} c_n \phi_n,$$

juntamente con la ecuación de Parseval, demuestran que puede considerarse $\mathscr{L}^2(\mu)$ como un espacio euclidiano de dimensión infinita (el llamado «espacio de Hilbert», en el que el punto f tiene coordenadas c_n, y las funciones ϕ_n son los vectores coordenadas.

EJERCICIOS

1. Si $f \geq 0$ y $\int_E f\, d\mu = 0$, demostrar que $f(x) = 0$ en casi todas partes en E. *Sugerencia*: Sea E_n el subconjunto de E en el cual $f(x) > 1/n$. Escribir $A = \bigcup E_n$. Entonces $\mu(A) = 0$ si y solo si $\mu(E_n) = 0$ para todo n.
2. Si $\int_A f\, d\mu = 0$ para todo subconjunto medible A de E, $f(x) = 0$ casi en todas partes en E.
3. Si $\{f_n\}$ es una sucesión de funciones medibles, probar que el conjunto de puntos x en los cuales $\{f_n(x)\}$ converge, es medible.
4. Si $f \in \mathscr{L}(\mu)$ en E y g es acotada y medible en E, $fg \in \mathscr{L}(\mu)$ en E.
5. Poniendo

$$g(x) = \begin{cases} 0 & (0 \leq x \leq \frac{1}{2}), \\ 1 & (\frac{1}{2} < x \leq 1), \end{cases}$$
$$f_{2k}(x) = g(x) \qquad (0 \leq x \leq 1),$$
$$f_{2k+1}(x) = g(1-x) \qquad (0 \leq x \leq 1).$$

es

$$\liminf_{n\to\infty} f_n(x) = 0 \qquad (0 \leq x \leq 1),$$

pero

$$\int_0^1 f_n(x)\, dx = \tfrac{1}{2}.$$

[Comparar con (77)].

6. Si

$$f_n(x) = \begin{cases} \dfrac{1}{n} & (|x| \leq n), \\ 0 & (|x| > n). \end{cases}$$

Entonces $f_n(x) \to 0$ uniformemente en R^1, pero

$$\int_{-\infty}^{\infty} f_n \, dx = 2 \qquad (n = 1, 2, 3, \ldots).$$

(Escribimos $\int_{-\infty}^{\infty}$ en lugar de $\int_{R^1}$). Así, la convergencia uniforme no implica la convergencia dominada en el sentido de Teorema 11.32. Sin embargo, en los conjuntos de medida finita, las sucesiones uniformemente convergentes de funciones acotadas. Satisfacen el Teorema 11.32.

7. Hallar una condición necesaria y suficiente para que $f \in \mathscr{R}(\alpha)$ en $[a,b]$. *Sugerencia*: Considerar el ejemplo 11.6(b) y el Teorema 11.33.

8. Si $f \in \mathscr{R}$ en $[a,b]$ y si $F(x) = \int_a^x f(t)\, dt$, $F'(x) = f(x)$ en casi todas partes en $[a,b]$.

9. Probar que la función F dada por (96) es continua en $[a,b]$.

10. Si $\mu(X) < +\infty$ y $f \in \mathscr{L}^2(\mu)$ en X, entonces $f \in \mathscr{L}(\mu)$ en X. Si

$$\mu(X) = +\infty,$$

esto es falso. Por ejemplo, si

$$f(x) = \frac{1}{1+|x|},$$

$f \in \mathscr{L}^2$ en R^1, pero $f \notin \mathscr{L}$ en R^1.

11. Si f, $g \in \mathscr{L}(\mu)$ en X, definir la distancia entre f y g por

$$\int_X |f-g| \, d\mu.$$

Probar que $\mathscr{L}(\mu)$ es un espacio métrico completo.

12. Suponer

(a) $|f(x,y)| \leq 1$ si $0 \leq x \leq 1$, $0 \leq y \leq 1$,

(b) para x fijo, $f(x,y)$ es función continua de y,

(c) para y fijo, $f(x,y)$ es función continua de x.

Poner

$$g(x) = \int_0^1 f(x, y)\, dy \qquad (0 \leq x \leq 1).$$

¿Es continua g?

13. Considerar las funciones

$$f_n(x) = \operatorname{sen} nx \qquad (n = 1, 2, 3, \ldots, -\pi \leq x \leq \pi)$$

como puntos de $\mathscr{L}^2$. Probar que el conjunto de estos puntos es cerrado y acotado, pero no compacto.

14. Probar que una función compleja es medible si y solo si $f^{-1}(V)$ es medible para todo conjunto abierto V en el plano.

15. Sea $\mathscr{R}$ el anillo de todos subconjuntos elementales de $(0, 1)]$. Si $0 < a \leq b \leq 1$, definir,

$$\phi([a, b]) = \phi([a, b)) = \phi((a, b]) = \phi((a, b)) = b - a,$$

pero siendo

$$\phi((0, b)) = \phi((0, b]) = 1 + b$$

si $0 < b \leq 1$. Demostrar que esto da una función de conjuntos aditiva ϕ en $\mathscr{R}$ que no es regular y que no puede ampliarse a una función de conjuntos aditiva numerable en un σ-anillo.

16. Suponer que $\{n_k\}$ es una sucesión creciente de enteros positivos y E es el conjunto de todos los $x \in (-\pi, \pi)$ en los que converge $\{\text{sen } n_k x\}$. Probar que $m(E) = 0$. *Sugerencia*: Para todo $A \subset E$,

$$\int_A \text{sen } n_k x \, dx \to 0,$$

y

$$2\int_A (\text{sen } n_k x)^2 \, dx = \int_A (1 - \cos 2n_k x) \, dx \to m(A) \quad \text{cuando } k \to \infty.$$

17. Suponer que $E \subset (-\pi, \pi)$; $m(E) > 0$, $\delta > 0$. Utilizar la desigualdad de Bessel para probar que hay a lo sumo un número finito de enteros n tales que sen $nx \geq \delta$ para todo $x \in E$.

18. Suponiendo que $f \in \mathscr{L}^2(\mu)$, $g \in \mathscr{L}^2(\mu)$, probar que

$$\left| \int f\bar{g} \, d\mu \right|^2 = \int |f|^2 \, d\mu \int |g|^2 \, d\mu$$

si y solo si hay una constante c tal que $g(x) = cf(x)$ en casi todas partes. (Comparar con el Teorema 11.35).

BIBLIOGRAFÍA

ARTIN, E.: "The Gamma Function," Holt, Rinehart and Winston, Inc., Nueva York, 1964.

BOAS, R. P.: "A Primer of Real Functions," Carus Mathematical Monograph No. 13, John Wiley & Sons, Inc., Nueva York, 1960.

BUCK, R. C. (ed.): "Studies in Modern Analysis," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

———: "Advanced Calculus," 2a. ed., McGraw-Hill Book Company, Nueva York, 1965.

BURKILL, J. C.: "The Lebesgue Integral," Cambridge University Press, Nueva York, 1951.

DIEUDONNE, J.: "Foundations of Modern Analysis," Academic Press, Inc., Nueva York, 1960.

FLEMING, W. H.: "Functions of Several Variables," Addison-Wesley Publishing Company, Inc., Reading, Mass., 1965.

GRAVES, L. M.: "The Theory of Functions of Real Variables," 2a. ed., McGraw-Hill Book Company, Nueva York, 1956.

HALMOS, P. R.: "Measure Theory," D. Van Nostrand Company, Inc., Princeton, N.J., 1950.

———: "Finite-dimensional Vector Spaces," 2a. ed., D. Van Nostrand Company, Inc., Princeton, N.J. 1958.

HARDY, G. H.: "Pure Mathematics," 9a. ed., Cambridge University Press, Nueva York, 1947.

———y ROGOSINSKI, W.: "Fourier Series," 2a. ed., Cambridge University Press, Nueva York, 1950.

HERSTEIN, I. N.: "Topics in Algebra," Blaisdell Publishing Company, Nueva York, 1964.

HEWITT, E., y STROMBERG, K.: "Real and Abstract Analysis," Springer Publishing Co., Inc., Nueva York, 1965.

KELLOGG, O. D.: "Foundations of Potential Theory," Frederick Ungar Publishing Co., Nueva York, 1940.

KNOPP, K.: "Theory and Application of Infinite Series," Blackie & Son, Ltd., Glasgow, 1928.

LANDAU, E. G. H.: "Foundations of Analysis," Chelsea Publishing Company, Nueva York, 1951.

MCSHANE, E. J.: "Integration," Princeton University Press, Princeton, N.J., 1944.

NIVEN, I. M.: "Irrational Numbers," Carus Mathematical Monograph No. 11, John Wiley & Sons, Inc., Nueva York, 1956.

ROYDEN, H. L.: "Real Analysis," The Macmillan Company, Nueva York, 1963.

RUDING, W.: "Real and Complex Analysis," 2a. ed., McGraw-Hill Book Company, Nueva York, 1974.

SIMMONS, G. F.: "Topology and Modern Analysis," McGraw-Hill Book Company, Nueva York, 1963.

SINGER, I. M., y THORPE, J. A.: "Lecture Notes on Elementary Topology and Geometry," Scott, Foresman and Company, Glenview, Ill., 1967.

SMITH, K. T.: "Primer of Modern Analysis," Bogden and Quigley, Tarrytown-on Hudson, N.Y., 1971.

SPIVAK, M.: "Calculus on Manifolds," W. A. Benjamin, Inc. Nueva York, 1965.

THURSTON, H. A.: "The Number System," Blackie & Son, Ltd., Londres-Glasgow, 1956.

LISTA DE SÍMBOLOS ESPECIALES

Los símbolos enumerados a continuación están acompañados por su significado breve y el número de la página en la que están definidos.

$\in$ pertenece a 3
$\notin$ no pertenece a 3
$\subset$, $\supset$ signos de inclusión 3
Q campo racional 3
$<$, $\leq$, $>$, $\geq$ signos de desigualdad . 3
sup mínima cota superior 4
inf máxima cota inferior 4
R campo real 9
$+\infty$, $-\infty$ infinitos 12, 29
$\bar{z}$ complejo conjugado 15
Re (z) parte real 15
Im (z) parte imaginaria 15
$|z|$ valor absoluto 15
Σ símbolo de suma 16, 62
R^k k-espacio euclidiano 17
$\mathbf{0}$ vector nulo 17
$\mathbf{x} \cdot \mathbf{y}$ producto interno 17
$|\mathbf{x}|$ norma de un vector 17
$\{x_n\}$ sucesión 28
$\bigcup$, $\cup$ unión 29
$\bigcap$, $\cap$ intersección 29
(a, b) segmento 33
$[a, b]$ intervalo 33
E^c complemento de 34
E' puntos límite de 37
$\bar{E}$ cerradura de 37
lím límite 50
$\rightarrow$ converge hacia 50, 104
lím sup límite superior 59
lím inf límite inferior 59
$g \circ f$ composición 92
$f(x+)$ límite por la derecha 101
$f(x-)$ límite por la izquierda 101
f', $\mathbf{f}'(\mathbf{x})$ derivadas 110, 119
$U(P,f)$, $U(P,f,\alpha)$, $L(P,f)L(P,f,\alpha)$
sumas de Riemann 130, 131

$\mathscr{R}$, $\mathscr{R}(\alpha)$ clases de funciones integrables de Riemann (Stieltjes) 130, 131
$\mathscr{C}(X)$ espacio de funciones continuas 161
$\| \ \|$ norma 150, 161, 352
exp función exponencial 192
D_N núcleo de Dirichlet 203
$\Gamma(x)$ función gama 207
$\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ base estándar 220
$L(X)$, $L(X, Y)$ espacios de transformaciones lineales 223
$[A]$ matriz 225
$D_j f$ derivada parcial 231
∇f gradiente 234
$\mathscr{C}'$, $\mathscr{C}''$ clases de funciones diferenciables 236, 254
det [ ] determinante 251
$J_f(\mathbf{x})$ Jacobiano 253
$\dfrac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}$ Jacobiano 254
I^k k-celda 265
Q^k k-símplex 267
dx_I k-forma básica 278
$\wedge$ símbolo de multiplicación 275
d operador diferenciación 282
ω_T transformada de 284
∂ operador frontera 291
$\nabla \times \mathbf{F}$ rotacional 305
$\nabla \cdot \mathbf{F}$ divergencia 305
$\mathscr{E}$ anillo de conjuntos elementales 327
m medida de Lebesgue 327, 333
μ medida 327, 333
$\mathfrak{M}_F$, $\mathfrak{M}$ familias de conjuntos medibles 330
$\{x \mid P\}$ conjunto con la propiedad . . 335
f^+, f^- parte positiva (negativa) de . 337
K_E función característica 338
$\mathscr{L}$, $\mathscr{L}(\mu)$, $\mathscr{L}^2$, $\mathscr{L}^2(\mu)$ clases de funciones integrables según Lebesgue 340, 352

ÍNDICE ANALÍTICO

Abel, N.H., 80, 186
Acotabilidad uniforme, 166
Adición,
fórmula de la, 191
(*véase también* Suma)
Aditividad, 325
numerable, 325
Álgebra, 173
auto-adjunta, 177
uniformemente cerrada, 173
(*véase también* Anillo)
Ángulo sólido, 319
Anillo, 325
Aplicación (*véase* Mapeo)
Aproximación cuadrática
media, 202
Arco, 146
Artin, E., 206, 210
Axiomas, 5, 6
de campo, 5
de cuerpo (*véase* Axiomas,
de campo)

Baire, teorema de, 49, 87
Base, 48
de un espacio vectorial, 220
estándar, 220
numerable, 49
Bellman, R., 213
Bessel, desigualdad de, 202, 354
Bohr-Mollerup, teorema de, 207
Bola, 34
Borel, conjunto de, 334
Brouwer, teorema de, 218
Buck, R.C., 210

Cadena, 290
afín, 290
diferenciable, 292
Cambio de variable, 142,
272, 284
Campo,
complejo, 13, 198
ordenado, 8, 22
conjunto, 3, 19, 23
k-tupla, 17
par, 13
real, 9
vectorial, 304
Cantor, G., 23, 32, 200
Cantor, conjunto de, 45, 87,
149, 180, 335
Casi en todas partes, 343
Cauchy,
criterio de, 57, 63, 157
prueba de condensación
de, 65
sucesión de, 23, 55, 88, 355
Celda, 33
Cerradura, 37
uniforme, 161, 173
Círculo de convergencia, 74
Clausura (*véase* Cerradura)
Colección, 29
Combinación lineal, 219
Complemento, 34
Compleción (*véase* Completez)
Completez, 88
Componente,
de una función, 94, 231
tangencial, 311
Composición, 92, 112, 136, 223
Conjugado, 15
Conjunto(s), 3
abierto, 34
acotado, 34
superiormente, 34
ajenos, 29
de Borel, 334
de Cantor, 44, 87, 149,
180, 335
cero, 105, 125
cerrado, 34
compacto, 39
conexo, 46
convexo, 34
denso, 10, 34
dependiente, 220
elemental, 327
finito, 27
independiente, 220
infinito, 27
a lo más numerable, 27
medible, 330, 335
no numerable, 27, 32, 44
no vacío, 3
numerable, 27
ordenado, 3
ortogonal de funciones, 200
ortonormal, 201, 354, 358
completo, 358
perfecto, 34, 44
relativamente abierto, 38
separados, 46
vacío, 3
Continuidad, 91
uniforme, 97
Contracción, 237
Contractiva, aplicación (*véase*
Contracción)
Convergencia, 50
absoluta, 76
de una integral, 149
acotada, 348
dominada, 347
de una integral, 149
puntual, 153

radio de, 74, 84
de series, 63
de sucesiones, 51
uniforme, 157
Coordenadas, 17, 220
Correspondencia uno a uno
(o inyectiva), 27
Cortadura, 18
Corte (*véase* Cortadura)
Cota,
inferior, 4
superior, 4
Criterio,
del cociente (*véase* Prueba,
de la razón)
de la raíz (*véase* Prueba,
de la raíz)
Cubierta, 39
abierta, 39
Cubo unidad, 267
Cuerpo complejo (*véase* Campo,
complejo)
Cuerpo de los reales (*véase*
Campo, real)
Cunningham, F., 179
Curva, 146
cerrada, 146
continuamente
diferenciable, 147
que llena un espacio, 180
rectificable, 146

David, P.J., 206
Decimales, 12
Dedekind, R., 23
Derivada, 110
direccional, 234
de una O-forma, 281
de una función vectorial, 119
integración de, 145, 351
de una integral, 143, 256, 351
normal, 322
de orden superior, 118
parcial, 231
de series de potencias, 185
total, 230
de una transformación, 230
Desigualdad del triángulo, 15,
17, 33, 150
Determinante, 250
de un operador, 253
producto de, 252
Diámetro, 56
Diferencia simétrica, 330
Diferenciación (*véase* Derivada)
Diferencial, 230
Dimensión, 220
Dirichlet, núcleo de, 203
Discontinuidad(es), 100
simple, 101
Distancia, 33
Divergencia, 305
Dominio, 26
de parámetros, 274

Eberlein, W.F., 198
Ecuación diferencial, 127, 182
Elemento de área, 307
En, 26
Entorno, (*véase* Vecindad)
Envolvente (*véase* Cubierta)
Equicontinuidad, 167
Equivalencia, 303
Esfera unitaria, 295, 301, 319
Espacio,
conexo, 46
euclidiano, 17, 33
de funciones continuas, 160
de funciones integrables,
340, 352
de Hilbert, 359
medible, 335
de medida, 335
métrico, 33
compacto, 39
completo, 58, 87, 161, 356
normal, 108
nulo, 246
separable, 48
vectorial, 17, 219
Euler, constante de, 212
Extensión, 106

Familia, 29
Fatou, teorema de, 347
Fejér,
núcleo de, 214
teorema de, 214
Fine, N.J., 107
Fleming, W. H., 304
Flip, 269
Forma, 275
básica, 277
cerrada, 298
de clase, 275
derivada de una, 282
exacta, 298
producto de, 279, 281
suma de, 276
Fourier, J.B., 200
coeficientes de, 200, 201
series de, 200, 201, 354
Frontera, 291
Función(es), 26
acotada, 96
analítica, 184
armónica, 322
Beta, 208
característica, 338
componente de, 94
de conjunto, 325
anillo, 325
regular, 327
constante, 91
continua, 91
por la derecha, 104
espacio de, 160
por la izquierda, 104
continuamente
diferenciable, 235
en ninguna parte, 164
convexa, 108
coordenada, 94
creciente, 102
decreciente, 102
diferenciable, 111, 217
en ninguna parte, 164
escalón, 139
exponencial, 191
Gramma, 206
integrables, espacios de,
340, 352
integrables según
Lebesgue, 340
integrables según
Riemann, 130
inversa, 96
inyectiva (*véase* Función, uno
a uno)
límite, 153
lineal, 221
logarítmica, 193
medible, 315
medible según Borel, 338
monótona, 102, 326
ortogonal, 200
periódica, 196, 205
producto de, 91
racional, 95
simple, 338
suma de, 91
sumable, 340
trigonométricas, 195
uniformemente continua, 97
uniformemente
diferenciable, 122
uno a uno, 27
valor absoluto, 94
vectorial, 91
derivada, de, 119
zeta, 152

Generador, 219
Gradiente, 234, 305
Gráfica, 106
Green,
identidades de, 322
teorema de, 274, 276,
295, 306

Havin, V.P., 121
Helly, teorema de selección
de, 179
Herstein, I.N., 70
Hewitt, E., 23
Hilbert, espacio de, 359
Hölder, desigualdad de, 150

Imagen, 26
inversa, 26,
Índice,
creciente, 277
de una curva, 216
Ínfimo, 4
Infinito, 12
Integración,
de una derivada, 149, 351
por partes, 145, 149, 151

Integral,
aditividad numerable de, 341
diferenciación de, 143, 256, 351
impropia, 149
inferior, 130, 131
de Lebesgue, 340
de línea, 275
de Riemann, 130
de Stieltjes, 131
superior, 130, 131
Interior, 47
Intersección, 29
Intervalo, 33, 327
del parámetro, 146
semiabierto, 33
Inversa de un operador lineal, 222
Isometría, 88, 182
Isomorfismo, 23

Jacobiano, 253

Kellogg, O.D., 305
Kestelman, H., 179
Knopp, K., 23, 67

Landau, E.G.H., 23
Laplaciano, 322
Lebesgue, H.L., 200
Lebesgue,
función integrable según, 340
integral de, 340
medida de, 334
teorema de, 166, 179, 344, 347
Leibnitz, G.W., 76
Ley,
anticonmutativa, 277
asociativa, 6, 30, 280
conmutativa, 6, 30
distributiva, 6, 22, 30
L'Hospital, regla de, 116, 121
Límite, 90, 104, 153
por la derecha, 100
inferior, 59
por la izquierda, 57
puntual, 153
subsecuencial, 54
superior, 59
Logaritmo, 24, 193
Longitud, 147

Mapeo, 26
abierto, 107, 241
afín, 288
continuamente diferenciable, 235
continuo, 91
inverso, 90
lineal, 221
localmente uno a uno, 241
primitivo, 268
uniformemente continuo, 97
(*véase también* Función)

Matriz, 225
columna, 233
producto, 226
renglón, 233
Máxima cota inferior, 4
Máximo, 96
local, 114
McShane, E.J., 338
Media aritmética, 85, 214
Medida, 334
cero, conjunto de, 334, 343
exterior, 328
de Lebesgue, 334
Mertens, F., 79
Mínima cota superior, 4
propiedad de, 5, 19
Mínimo, 96
Möbius, cinta de, 322
Módulo (*véase* Valor absoluto)
Multiplicación (*véase* Producto)

Newton, método de, 126
Nijenhuis, A., 241
Niven, I., 70, 213
Norma, 18, 150, 161, 352
de un operador, 223
supremum, 161
Número(s),
algebraico, 47
de arrollamiento, 216
cardinal, 27
complejo, 13
decimal, 12
finito, 13
irracional, 1, 10, 69
negativo, 8
no negativo, 64
positivo, 8, 9
racional, 1
real, 9

Operador,
identidad, 251
lineal, 222
Orden, 3, 17
lexicográfico, 24
Orientación,
negativa, 289
positiva, 289
Origen, 17

Parseval, teorema de, 205, 213, 354, 358
Parte,
imaginaria, 15
real, 15
Partición, 129
de la unidad, 271
Plano, 18
complejo, 18
tangente, 308
Poincaré, lema de, 298, 304
Polinomio, 94
trigonométrico, 199
Primos, 212
Proceso diagonal, 32, 168

Producto, 6
de Cauchy, 77
de determinantes, 252
de elementos de un campo, 6
escalar, 17
de formas, 279, 281
de funciones, 91
interior, 17
de matrices, 226
de números complejos, 13
de números reales, 21, 22
de series, 77
de transformaciones, 222
Problema de valor inicial, 127, 182
Propiedad reflexiva, 27
Proyección, 246
Prueba,
de comparación, 64
de la integral, 149
de la raíz, 70
de la razón o del cociente, 70
Punto,
de acumulación (*véase* Punto, de condensación)
aislado, 34
de condensación, 49
fijo, 125
teoremas de, 125, 218, 237
interior, 34
límite, 34
silla, 259

Radio, 34,
de convergencia, 74, 84
Raíz, 10
cuadrada, 2, 86, 126
Rango,
de una función, 24, 222
de una transformación lineal, 246
Recta, 18
Recta real, 18
Recubrimiento (*véase* Cubierta)
Refinamiento, 132
común, 132
Regla de la cadena, 112, 230
Relación de equivalencia, 27
Reordenamiento, 80
Reordenación (*véase* Reordenamiento)
Representación, estándar, 278
Residuo, 227, 263
Restricción, 106
Riemann, B., 81, 200
integral de, 130
Riemann-Stieltjes, integral de, 131
Riesz-Fischer, teorema de, 357
Robison, G.B., 198
Rotacional, 305

Schoenberg, I.J., 180
Schwarz, desigualdad de, 16, 150, 353
Segmento, 34

Separación de puntos, 173
Serie(s), 63
absolutamente convergente, 76
alternante, 76
binomial, 216
convergente, 63
divergente, 63
geométrica, 64
infinita, 63
no-absolutamente convergente, 76
de potencias, 73, 184
producto de, 77
trigonométricas, 200
uniformemente convergente, 168
Símplex, 267
afín, 288
diferenciable, 292
estándar, 288
orientado, 289
o-anillo, 325
Singer, I.M., 304
Sistema extendido de los números reales, 12
Sobre, 26
Soporte, 266
Spivak, M., 296, 304
Stark, E.L., 214
Stieltjes, integral de, 131
Stirling, fórmula de, 214-215
Stone-Weierstrass, teorema de, 174, 204, 266
Stokes, teorema de, 274, 295, 312
Stromberge, K., 23
Subaditividad, 329
Subcampo, 9, 15
Subconjunto, 3
denso, 10, 34
propio, 3
Subcubierta, 39
Subsucesión, 55
Sucesión, 28
acotada, 51
puntualmente, 166
de Cauchy, 55, 88, 355
convergente, 50
creciente, 58
divergente, 50
doble, 154
de funciones, 153
monótona, 58
puntualmente convergente, 153
uniformemente acotada, 166
uniformemente convergente, 168
Suma, 6
de elementos de un campo, 6
de formas, 277
de funciones, 91
de números complejos, 13
de números reales, 19
parcial, 63, 200
de series, 63
de símplexes orientados, 290
de transformaciones lineales, 223
de vectores, 17
Sumación por partes, 74
Superficie, 274
Supremum, 161

Taylor,
polinomio de, 264
teorema de, 118, 123, 189, 263
Teorema,
de la convergencia dominada, 166, 179, 347
de la convergencia monótona, 344
de la divergencia, 274, 295, 312
de existencia, 182
de la función implícita, 242
de la función inversa, 238
fundamental del cálculo, 144, 350
de Heine-Borel, 43
de localización, 204
del rango,
de unicidad, 127, 279
del valor medio, 115, 254
Thorpe, J.A., 304
Thurston, H.A., 23
Toro, 259, 260, 309
Transformación,
invertible, 222
lineal, 221
(*véase también* Función; Mapeo)
Transitividad, 27

Unión, 29

Valor, 26
absoluto, 15
intermedio, 100, 107, 115
Variable de integración, 131
Vecindad, 34
Vector, 17
columna, 226
normal, 308
nulo, 17
tangente, 310
unitario, 234
Volumen, 276, 306

Weierstrass,
prueba de, 147
teorema de, 44, 170